

**Universidade Técnica de Lisboa**

Instituto Superior de Economia e Gestão

Mestrado em Ciências Actuarias

**Aplicação da Teoria de Valores Extremos  
à Actividade Seguradora**

Jeferino Manuel dos Santos

Orientação

Prof. Doutora Maria de Lourdes Caraças Centeno

Júri

Prof. Doutora Maria de Lourdes Caraças Centeno

Prof. Doutora Maria Isabel Fraga Alves

Prof. Doutor Alfredo Duarte Egídio dos Reis

Outubro 2003



## RESUMO

---

O objectivo principal deste trabalho é realçar a importância da Teoria de Valores Extremos na actividade seguradora. São apresentados de uma forma sucinta alguns dos principais resultados ligados a esta teoria. São apresentadas algumas estatísticas que possibilitam a simplificação do processo de reconhecimento de dados de cauda pesada. A modelação da cauda é um assunto de particular interesse, são apresentados dois métodos de modelação da cauda, um pelo ajustamento de uma distribuição de Pareto Generalizada, outro pela aplicação de um método semi-paramétrico adaptativo. No fim, os resultados obtidos por cada um dos modelos são integrados como módulo num modelo de solvência.

**Palavras chave:** Teoria de Valores Extremos, dados de cauda pesada, Distribuição de Pareto Generalizada, estimação semi-paramétrica adaptativa, indemnizações agregadas, modelo de solvência.

## ABSTRACT

---

The main purpose of this dissertation is to enhance the importance of Extreme Value Theory in the insurance sector. A short introduction to the main results inherent in this theory is presented. Also, a set of statistics to simplify the recognition process of heavy tailed data is provided. Tail modelling is a subject of particular interest in this dissertation, two approaches are presented, one by fitting a Generalized Pareto Distribution, other by modelling by means of a semi-parametric adaptive method. In the last part, the results of these approaches are integrated as a module in a broader solvency model.

**Keywords:** Extreme Value Theory, heavy tailed data, Generalized Pareto Distribution, semi-parametric adaptive estimation, aggregate loss, solvency model.

# Índice

<b>Prefácio</b>	<b>9</b>
<b>Agradecimentos</b>	<b>11</b>
<b>Capítulo 1 – Teoria de Valores Extremos</b>	<b>13</b>
1.1. Alguns Resultados Preliminares	13
1.2. Caracterização dos Domínio de Atracção	16
1.3. Distribuição de Pareto Generalizada	27
<b>Capítulo 2 – Análise e Apresentação dos Dados</b>	<b>31</b>
<b>Capítulo 3 – Modelação dos dados</b>	<b>37</b>
3.1 Modelação pela Distribuição de Pareto Generalizada	37
3.2 Abordagem semi paramétrica	44
3.2.1 Estimação Sob Condição de Domínio de Atracção para Máximos	44
3.2.2 Determinação do Nível Óptimo de $k$	48
3.3 Discussão de diferentes métodos	58
<b>Capítulo 4 – Teste de Solvência</b>	<b>61</b>
4.1 Teste da Solvência	61
4.2 Modelação das Indemnizações Agregadas	64
4.3 Discussão	72
<b>Capítulo 5 – Conclusão</b>	<b>77</b>
<b>Anexo I</b>	<b>79</b>
<b>Bibliografia</b>	<b>83</b>

## Lista de Figuras

**Figura 1.1.1** – Funções de densidade das formas possíveis de  $H(\cdot)$ , com  $\alpha = 1$  para a Fréchet e para a Weibull. [p. 16]

**Figura 1.3.1** - Funções de densidade das formas possíveis de  $G(\cdot)$ , para diferentes valores de  $\xi$ , com  $\beta = 1$  e  $\nu = 0$ . [p. 28]

**Figura 2.1** – Comportamento dos sinistros extremos verificados entre 1993 e 2001. [p. 32]

**Figura 2.2** – No gráfico da esquerda estão representadas a distribuição empírica e a distribuição de Pareto. O gráfico da direita tem a escala ampliada na cauda. [p. 33]

**Figura 2.3** – Comparação com os quantis da distribuição de Pareto. [p. 33]

**Figura 2.4** - Função de excesso médio empírico. [p. 34]

**Figura 2.5** -  $R_n(p)$  para varios valores de  $p$ . [p. 35]

**Figura 3.1.1** – Valores de  $\hat{\xi}$ , da DPG, para diferentes níveis de  $u$ . [p. 40]

**Figura 3.1.2** – Função de distribuição ajustada aos excessos com  $u = 550.000$  (à esquerda) e  $u = 600.000$  (à direita). [p. 41]

**Figura 3.1.3** – Função de distribuição ajustada aos excessos com  $u = 550.000$  (à esquerda) e  $u = 600.000$  (à direita), com a cauda ampliada. [p. 41]

**Figura 3.1.4** – Comparação dos quantis dos excessos com as distribuições ajustadas com  $u = 550.000$  (à esquerda) e  $u = 600.000$  (à direita). [p. 42]

**Figura 3.1.5** – No gráfico da esquerda estão representadas a distribuição empírica e a distribuição ajustada com  $u = 550.000$  e  $u = 600.000$ . O gráfico da direita tem a escala ampliada na cauda. [p. 43]

**Figura 3.1.6** – Comparação dos quantis com os da distribuição ajustada com  $u = 550.000$  e  $u = 600.000$ . [p. 44]

**Figura 3.2.1** – Horror Hill plot. [p. 49]

**Figura 3.2.2** – Resultados para diferentes valores de  $n_1$  dado  $\xi_{n,aux} = 0,765$ . [p. 53]

**Figura 3.2.3** – Resultados para diferentes valores de  $n_1$  dado diferentes valores de  $\xi_{n,aux}$ . [p. 54]

**Figura 3.2.4** – Resultados para diferentes valores de  $n_1$ , segundo a abordagem de Danielsson. [p. 56]

**Figura 3.2.5** – No gráfico da esquerda pode-se observar a distribuição empírica e a distribuição ajustada, segunda a abordagem de Danielsson. O gráfico da direita tem a escala ampliada na cauda. [p. 57]

**Figura 3.2.6** – Comparação dos quantis com os da distribuição ajustada segundo a abordagem de Danielsson. [p. 58]

**Figura 3.3.1** – Comparação dos resultados finais de diferentes estatísticas auxiliares. [p. 59]

**Figura 3.3.2** – Comparação dos quantis com os da distribuição ajustada em 3.1 com  $u = 550.000$  e a distribuição ajustada segundo a abordagem de Danielsson. [p. 60]

**Figura 4.1.1** – Economic Capital e a Solvência em termos esquemáticos. [p. 62]

**Figura 4.1.2** – Função de Densidade (à esquerda) e Função de Distribuição (à direita) da variação do Economic Capital. [p. 63]

**Figura 4.2.1** –  $R_{n_i}(p)$  das indenizações da classe  $A_i$  do Modelo 1 e Modelo 2. [p. 68]

**Figura 4.2.2** – Função de distribuição condicionada empírica das indenizações da classe  $A_i$  do Modelo 1 e Modelo 2. [p. 68]

**Figura 4.2.3** – Funções de densidade (à esquerda) e de distribuição (à direita) de  $S^{(1)}$ ,  $S^{(2)}$ ,  $\hat{S}^{(1)}$  e  $\hat{S}^{(2)}$ . [p. 70]

**Figura 4.2.4** – Comparação dos quantis estimados de  $S^{(1)}$ ,  $S^{(2)}$ ,  $\hat{S}^{(1)}$  e  $\hat{S}^{(2)}$ . [p. 71]

**Figura 4.2.5** – Comparação dos quantis das indenizações agregadas retidas. [p. 72]

**Figura 4.3.1** – Comparação dos quantis dos valores simulados de  $S_1^{(1)}$  e de  $S_1^{(2)}$  e os quantis dos métodos aproximados. [p. 75]



## PREFÁCIO

---

Pretende-se com o presente trabalho analisar as indemnizações relacionadas com acontecimentos extremos. São de particular interesse os acontecimentos catastróficos que conduzem a indemnizações de elevado montante, não se reduzindo a indemnizações associadas a fenómenos naturais ou desastres causados pelo Homem, mas também a indemnizações invulgares de acordo com o risco em causa.

Uma modelação adequada das indemnizações extremas é essencial para a actividade de uma Companhia de Seguros, dado que isso permitiria:

- definir um nível apropriado do prémio;
- modelar adequadamente o resseguro;
- determinar o nível do capital necessário por forma a minimizar a probabilidade de insolvência.

Uma modelação possível da severidade é o ajustamento de uma função de distribuição às indemnizações individuais, como em Hogg *et al.*(1984). Porém, nessa abordagem, as indemnizações de valor intermédio acabam por ter um peso determinante na

distribuição escolhida e no valor do parâmetro estimado, podendo conduzir a uma sub ou sobre estimação da cauda.

Em muitos casos, dependendo do risco em causa, é importante que se modele isoladamente a cauda da distribuição da severidade por forma a projectar, com maior precisão, os quantis elevados.

Um dos objectivos principais desta dissertação consiste em aplicar algumas das recentes abordagens associados à Teoria de Valores Extremos, por forma a obter estimadores adequados para a cauda.

No Capítulo 1 são apresentados, de forma sumária, alguns dos principais resultados da Teoria de Valores Extremos.

Os dados, objecto de análise nesta dissertação, são apresentados e o seu comportamento estudado no Capítulo 2.

No Capítulo 3 são abordados dois estimadores da cauda. Um baseado numa abordagem paramétrica, o outro, numa abordagem semi-paramétrica adaptativa.

Na primeira parte do Capítulo 4 são discutidas algumas ideias sobre os testes de solvência. Na segunda parte, é modelado uma parte de um modelo de solvência aplicando os resultados obtidos no Capítulo 3, nomeadamente, a parte relacionada com as indemnizações agregadas e as indemnizações agregadas retidas.

O Capítulo 5, é reservado para a conclusão.

## **AGRADECIMENTOS**

---

À Prof. Doutora Maria de Lourdes Centeno, do Instituto Superior de Economia e Gestão. A sua orientação e sugestões em muito contribuíram para uma melhor exposição de ideias e conceitos.

À Ms. Ulrike Leyherr, do Allianz Group. Pelo seu contributo no fornecimento de dados que foram objecto de análise.

À Prof. Doutora Ivette Gomes, da Faculdade de Ciências da Universidade de Lisboa pelos textos gentilmente facultados.

À Allianz Portugal, pelo estímulo e facilidades concedidos.

À Dra. Teresa Brantuas, responsável pela Direcção de Actuariado da Allianz Portugal, e a todos os membros que compõem a sua equipa de trabalho. O seu apoio e incentivo foram importantes.

A todos aqueles que, dedicadamente, comentaram, criticaram e sugeriram ideias e opiniões.

À minha família e aos meus amigos que acreditaram na conclusão deste trabalho.



## CAPÍTULO 1

---

### Teoria de Valores Extremos

Neste capítulo são apresentados alguns dos principais resultados da Teoria de Valores Extremos. Os resultados mais importantes, nomeadamente, o Teorema de Fisher – Tippett e as três distribuições de extremos são expostos na Secção 1.1 deste capítulo. Na Secção 1.2, são caracterizados os domínios de atracção. Seguidamente, na Secção 1.3, são discutidas algumas das propriedades da Distribuição de Pareto Generalizada.

#### 1.1. Alguns Resultados Preliminares

Seja  $X_1, X_2, \dots$  uma sucessão de v.a. i.i.d. com função de distribuição  $F$ . Seja  $M_n$  o máximo de uma amostra aleatória de dimensão  $n$ , i.e.,

$$M_n = \max(X_1, X_2, \dots, X_n), \quad n \geq 1.$$

Os resultados que se apresentam nos próximos parágrafos têm a ver essencialmente com o máximo da amostra, na medida em que, por um lado, o mínimo assume uma menor importância na actividade seguradora, por outro, a conversão dos resultados na maior parte das vezes é quase imediata, atendendo à seguinte igualdade:

$$\min(X_1, X_2, \dots, X_n) = -\max(-X_1, -X_2, \dots, -X_n).$$

A função de distribuição (f.d.) exacta de  $M_n$  é  $F^n(x)$ , visto que

$$\text{Prob}[M_n \leq x] = \text{Prob}\left\{\bigcap_{i=1}^n [X_i \leq x]\right\} = \prod_{i=1}^n \text{Prob}[X_i \leq x] = F^n(x), \quad x \in \mathfrak{R}, \quad n \in \mathfrak{N}.$$

Seja

$$x_f = \sup\{x \in \mathfrak{R} : F(x) < 1\} \leq \infty,$$

o limite superior do suporte da f.d.  $F$ . Então,

$$M_n \xrightarrow{q.c.} x_f, \quad n \rightarrow \infty, \quad (1.1)$$

dado que para  $x < x_f$

$$\text{Prob}[M_n \leq x] = F^n(x) \rightarrow 0, \quad n \rightarrow \infty.$$

Enquanto, para  $x \geq x_f$

$$\text{Prob}[M_n \leq x] = F^n(x) = 1.$$

Por conseguinte,  $M_n \xrightarrow{p} x_f$ , e, dado que,  $M_n$  é não decrescente em relação a  $n$ , a convergência é quase certa.

Contudo, este resultado não é muito expressivo. É importante conhecer a magnitude que o máximo de uma amostra de dimensão  $n$  pode assumir. Esta é a principal preocupação da Teoria de Valores Extremos.

Um dos objectivos desta teoria consiste no estudo do comportamento dos máximos e dos mínimos das amostras, procurando encontrar aproximações para as suas distribuições. Isto é, aproximações de  $F^n$  ou de  $\bar{F}^n$ , que desempenhem um papel

semelhante ao da distribuição Normal no Teorema do Limite Central (TLC), como aproximação da distribuição da média da amostra.

Para que tal aproximação seja possível, a expressão (1.1) sugere a necessidade em transformar  $M_n$  numa variável  $X$ , tal que

$$c_n^{-1}(M_n - d_n) \stackrel{d}{=} X.$$

O Teorema de Fisher –Tippett, também conhecido por Teorema dos Tipos Extremais, demonstrado em 1943 por Gnedenko, é um dos resultados mais importantes da Teoria de Valores Extremos.

**Teorema 1.1.1** (Teorema de Fisher –Tippett)

Seja  $X_1, X_2, \dots$  uma sucessão de v.a. i.i.d. com função de distribuição  $F$  e seja, para  $n \geq 1$ ,  $M_n = \max(X_1, X_2, \dots, X_n)$ .

Se existir uma sucessão  $c_n$  de termos positivos, uma sucessão real de  $d_n$  e uma f.d.

$H(\cdot)$  não degenerada tais que, para cada  $x$ ,

$$\text{Prob}[M_n \leq c_n x + d_n] = F^n(c_n x + d_n) \xrightarrow{n \rightarrow \infty} H(x),$$

então as únicas formas possíveis de  $H(\cdot)$  são:

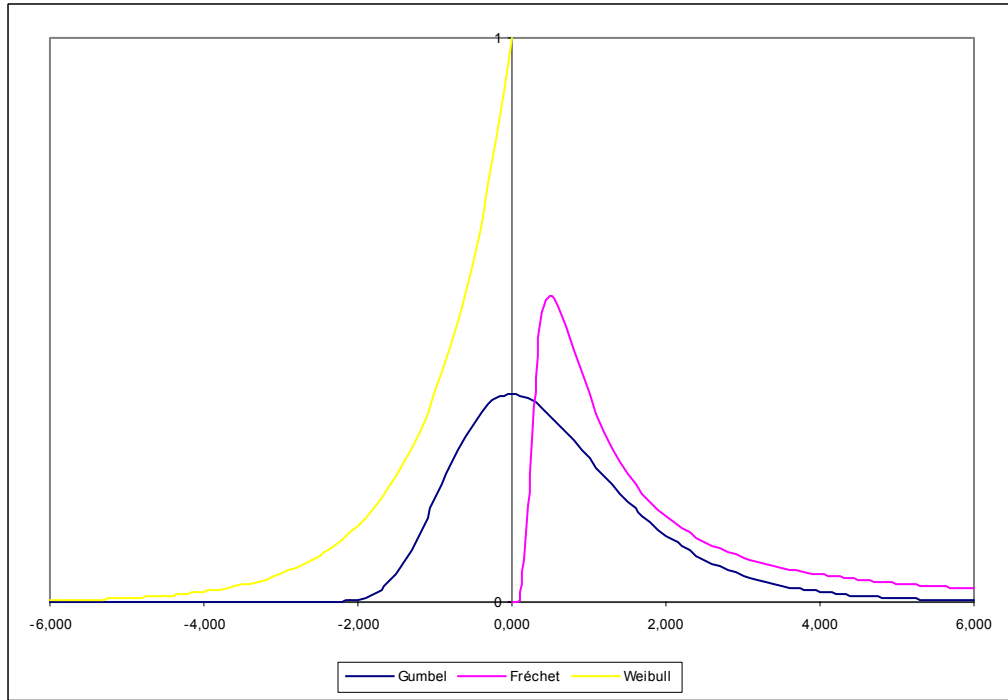
$$\text{Tipo I (Gumbel)} \quad H(x) \equiv \Lambda(x) = \exp\{-e^{-x}\}, \quad x \in \mathfrak{R};$$

$$\text{Tipo II (Fréchet)} \quad H(x) \equiv \Phi_\alpha(x) = \begin{cases} 0 & , x \leq 0 \\ \exp\{-x^{-\alpha}\} & , x > 0 \end{cases}, \quad \alpha > 0;$$

$$\text{Tipo III (Weibull)} \quad H(x) \equiv \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\} & x < 0, \quad \alpha > 0 \\ 1 & x \geq 0, \quad \alpha > 0 \end{cases}.$$

Dem. in Resnick (1987), p. 9-10.

No gráfico seguinte estão representadas as funções de densidade para as diferentes formas possíveis de  $H(\cdot)$ .



**Figura 1.1.1** – Funções de densidade das formas possíveis de  $H(\cdot)$ , com  $\alpha = 1$  para a Fréchet e para a Weibull.

Embora os três modelos se distingam entre si, existe uma relação matemática, pois:

$$X \sim \Phi_\alpha \Leftrightarrow \ln X^\alpha \sim \Lambda \Leftrightarrow -X^{-1} \sim \Psi_\alpha.$$

## 1.2. Caracterização dos Domínio de Atracção

Um dos tópicos mais importantes em Teoria de Valores Extremos é a caracterização dos domínios de atracção para máximos. Consiste no estabelecimento das condições necessárias e suficientes de  $F$  para pertencer a um determinado domínio de atracção e definir os valores apropriados de  $c_n$  e  $d_n$ .



**Definição 1.2.1** (Domínio de Atracção para Máximos)

$F$  pertence ao domínio de atracção de  $H$  para Máximos, e escreve-se  $F \in D(H)$ , se existir um par de sucessões  $c_n \geq 0$ ,  $d_n \in \mathfrak{R}$  tal que

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} H.$$

**Teorema 1.2.2** (Domínio de Atracção para Máximos)

$F$  pertence ao domínio de atracção de  $H$  para Máximos se e só se existir um par de sucessões  $c_n \geq 0$ ,  $d_n \in \mathfrak{R}$  tal que

$$\lim_{n \rightarrow \infty} n\bar{F}(c_n x + d_n) = -\ln H(x).$$

O conceito da *função de variação regular* é essencial para a caracterização do domínio de atracção das leis de Fréchet e de Weibull.

**Definição 1.2.3.** (Função de Variação Regular)

Uma função  $U$  mensurável,  $\mathfrak{R}^+ \rightarrow \mathfrak{R}^+$ , diz-se que é de variação regular no infinito com índice  $\rho$ , e escreve-se  $U \in RV_\rho$ , se para  $x > 0$ :

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\rho.$$

Uma função  $U(x)$  diz-se que é de variação regular em zero se  $U(1/x)$  for de variação regular no infinito.

**Definição 1.2.4.** (Função de Variação Lenta)

Uma função  $L(x)$  mensurável,  $\mathfrak{R}^+ \rightarrow \mathfrak{R}^+$ , diz-se que é de variação lenta se para  $x > 0$ :

$$\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1.$$

Que não é mais do que uma função de variação regular com índice zero.

Se  $U \in RV_\rho$  então  $U(x)/x^\rho \in RV_0$ . Definindo  $L(x) = U(x)/x^\rho$  torna-se possível representar  $U(x) = x^\rho L(x)$ . Desta forma, na maioria das situações, basta conhecer as propriedades das funções de variação lenta para se conhecer as propriedades das funções de variação regular.

**Teorema 1.2.5** (Propriedades das Funções de Variação Lenta)

Se  $L(x)$  é uma função de variação lenta:

- i. Para todo o  $\gamma > 0$ ,  $x^\gamma L(x) \rightarrow \infty$  e  $x^{-\gamma} L(x) \rightarrow 0$ , quando  $x \rightarrow \infty$ ;
- ii.  $\ln L(x)/\ln(x) \rightarrow 0$ , quando  $x \rightarrow \infty$ ;
- iii. A função  $L^\alpha(x)$ ,  $\alpha \in \mathbb{R}$ , é de variação lenta.

Se  $L_1(\cdot)$  e  $L_2(\cdot)$  forem funções de variação lenta,

- iv.  $L_1(\cdot) + L_2(\cdot)$  e  $L_1(\cdot)L_2(\cdot)$  são funções de variação lenta;
- v. Se  $L_2(\cdot) \rightarrow \infty$  quando  $x \rightarrow \infty$ , então  $L_1(L_2(\cdot))$  é de variação lenta.

**Definição 1.2.6** (Função Inversa)

Seja  $H$  uma função não decrescente em  $\mathbb{R}$ . Define-se a função inversa (contínua à esquerda) de  $H$  como

$$H^\leftarrow(y) = \inf \{s \in \mathbb{R} : H(s) \geq y\}.$$

(Assumindo que o ínfimo de um conjunto vazio é  $\infty$ .)

**Teorema 1.2.7** (Domínio de atracção de  $\Phi_\alpha$ )

Uma f.d.  $F$  pertence ao domínio de atracção para máximos de  $\Phi_\alpha$ ,  $\alpha > 0$ , se e só se

$$\bar{F}(x) = x^{-\alpha} L(x), \quad L \in RV_0.$$

Se  $F \in D(\Phi_\alpha)$ , então

$$c_n^{-1} M_n \xrightarrow{d} \Phi_\alpha,$$

onde  $c_n = F^{\leftarrow}(1 - n^{-1})$ .

Dem. in Resnick (1987), p. 54-57.

A lei de Fréchet é domínio de atracção de um conjunto de distribuições classificadas de cauda pesada, no sentido de  $E(X^\delta) = \infty$  para  $\delta > \alpha$ . A log-gamma, a Pareto ou a Burr, são alguns dos exemplos de distribuições que pertencem a este domínio de atracção. Tratam-se de distribuições que possuem algum interesse na modelação do valor das indemnizações em alguns ramos da actividade seguradora.

**Teorema 1.2.8** (Domínio de atracção de  $\Psi_\alpha$ )

Uma f.d.  $F$  pertence ao domínio de atracção para máximos de  $\Psi_\alpha$ ,  $\alpha > 0$ , se e só se

$$x_f < \infty \text{ e } \bar{F}(x_f - x^{-1}) = x^{-\alpha} L(x), \quad L \in RV_0.$$

Se  $F \in D(\Psi_\alpha)$ , então

$$c_n^{-1} (M_n - d_n) \xrightarrow{d} \Psi_\alpha,$$

onde  $c_n = x_f - \bar{F}^{\leftarrow}(1 - n^{-1})$  e  $d_n = x_f$ .

Dem. in Resnick (1987), p. 59-62.

A lei de Weibull é domínio de atracção para máximos de algumas distribuições que têm limitado o lado direito do suporte, tais como, a Uniforme, a Beta, entre outras. São distribuições que assumem menor importância na actividade seguradora devido ao limite superior do suporte ser finito.

Para caracterizar o domínio de atracção para máximos da Gumbel, é necessário introduzir alguns conceitos adicionais.

**Definição 1.2.9** (Função de Variação Rápida)

Uma função  $h$  mensurável,  $\mathfrak{R}^+ \rightarrow \mathfrak{R}^+$ , diz-se que é de variação rápida, e escreve-se

$h \in RV_{-\infty}$ , se

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = \begin{cases} 0 & \text{se } t > 1, \\ \infty & \text{se } 0 < t < 1. \end{cases}$$

**Teorema 1.2.10** (Propriedades das Funções de Variação Rápida)

i. Seja  $h \in RV_{-\infty}$ , não crescente, para  $z > 0$  e  $\alpha \in \mathfrak{R}$

$$\int_z^\infty t^\alpha h(t) dt < \infty$$

e

$$\lim_{x \rightarrow \infty} \frac{x^{\alpha+1} h(x)}{\int_x^\infty t^\alpha h(t) dt} = \infty;$$

Para  $z = 1$ , o inverso também é válido.

ii. Se  $h \in RV_{-\infty}$ , existem as funções  $c$  e  $\delta$  tal que  $c(x) \rightarrow c_0 \in (0, \infty)$ ,

$$\lim_{x \uparrow \infty} \delta(x) = -\infty \text{ e para } z > 0,$$

$$h(x) = c(x) \exp \left\{ \int_z^x \frac{\delta(u)}{u} du \right\}, \quad x \geq z.$$

O inverso também é válido.

**Definição 1.2.11** (Função de variação- $\Gamma$ )

Uma função  $U$ , não decrescente, diz-se de variação- $\Gamma$ , e escreve-se  $U \in \Gamma$ , se estiver definida num intervalo  $(a, b)$ , se  $\lim_{x \uparrow b} U(x) = \infty$ , e se existir uma função  $f$  definida em  $(a, b)$  tal que, para qualquer  $x$ ,

$$\lim_{t \rightarrow b} \frac{U(t + xf(t))}{U(t)} = e^x.$$

A função  $f$  chama-se função auxiliar e é assintoticamente única.

**Definição 1.2.12** (Função de variação- $\Pi$ )

Uma função  $V$ , não negativa e não decrescente definida num intervalo semi-infinito  $(z, \infty)$  diz-se de variação- $\Pi$ , e escreve-se  $V \in \Pi$ , se existirem funções  $a(t) > 0$ ,  $b(t) \in \Re$  tal que para  $x > 0$

$$\lim_{t \rightarrow \infty} \frac{V(tx) - b(t)}{a(t)} = \log x.$$

A função  $a(t)$  é assintoticamente única.

**Teorema 1.2.13** (Domínio de atracção de  $\Lambda$ )

Uma f.d.  $F$  pertence ao domínio de atracção para máximos de  $\Lambda$ , se e só se

$$\lim_{x \rightarrow x_f} \frac{(1 - F(x)) \int_x^{x_f} \int_y^{x_f} (1 - F(t)) dt dy}{\left( \int_x^{x_f} (1 - F(t)) dt \right)^2} = 1.$$

Todos os integrais envolvidos são finitos. Nesse caso  $1/(1-F(x)) \in \Gamma$  e duas escolhas possíveis para a função auxiliar  $f$

$$f(t) = \frac{\int_x^{x_f} \int_y^{x_f} (1-F(t)) \, dt dy}{\int_x^{x_f} (1-F(t)) \, dt}$$

ou

$$f(t) = \frac{\int_x^{x_f} (1-F(t)) \, dt}{(1-F(x))}$$

e

$$d_n = (1/(1-F))^\leftarrow(n)$$

$$c_n = f(d_n)$$

são escolhas aceitáveis.

*Dem. in Resnick (1987), p. 48-50.*

A lei de Gumbel é domínio de atracção de um vasto conjunto de distribuições, onde inclui as distribuições de cauda “normal”, ou moderadamente pesada, e as distribuições com o lado direito do suporte limitado. São disso exemplo a Normal, a Log-normal (que não é considerado pesada), a Gamma, a Benktander-I e a Benktander-II. Ao contrário das leis de Frechét e de Weibull, onde as constantes  $c_n$  e  $d_n$  podem ser obtidas de forma quase imediata, pelo Teorema 1.2.13, é possível verificar, que no caso da Gumbel, existe uma multiplicidade  $c_n$  e  $d_n$ . Isso deve-se à enorme variedade de distribuições que pertencem ao seu domínio de atracção.

As funções de distribuição que têm como domínio de atracção a lei de Gumbel, possuem uma propriedade importante em relação aos momentos.

**Corolário 1.2.14** (Existência de Momentos)

Dada uma v.a.  $X$ , com f.d.  $F \in D(\Lambda)$  e com  $x_f = \infty$ , então  $\bar{F} \in RV_{-\infty}$ . Em particular,

$$E[(X^+)^{\alpha}] < \infty \text{ para todo o } \alpha > 0, \text{ onde } X^+ = \max(0, X).$$

A segunda parte pode ser verificada pelo Teorema 1.2.10 i.

O cálculo de  $c_n$  e  $d_n$  é simplificado com o conceito que se segue.

**Definição 1.2.15** (Equivalência de Cauda)

Duas f.d.s  $F$  e  $G$  dizem-se que têm caudas equivalentes, se tiverem o mesmo limite superior do suporte, i.e., se  $x_f = x_g$  e  $\lim_{x \uparrow x_f} \bar{F}(x)/\bar{G}(x) = c$ ,  $0 < c < \infty$ .

**Teorema 1.2.16**

Dadas duas f.d.s  $F$  e  $G$ , seja  $H_1$  e  $H_2$  duas distribuições de valor extremo. Seja

$F \in D(H_1)$  e que  $F^n(c_n x + d_n) \rightarrow H_1(x)$ , com  $c_n > 0$  e  $d_n \geq 1$ . Então

$G^n(c_n x + d_n) \rightarrow H_2(x)$  se e só se para algum  $c > 0$ ,  $d \in \mathfrak{R}$

$H_2 = H_1(cx + d)$ ,  $F$  e  $G$  tiverem caudas equivalentes com suporte superior  $x_f$  e se

i.  $H_1 = \Phi_{\alpha}$ , então  $d = 0$  e  $\lim_{x \uparrow \infty} (1 - F(x))/(1 - G(x)) = c^{\alpha}$ ;

ii.  $H_1 = \Psi_{\alpha}$ , então  $d = 0$  e  $\lim_{x \uparrow x_f} (1 - F(x))/(1 - G(x)) = c^{-\alpha}$ ;

iii.  $H_1 = \Lambda$ , então  $c = 1$  e  $\lim_{x \uparrow x_f} (1 - F(x))/(1 - G(x)) = e^d$ .

Dem. in Resnick (1987), p. 68-71.

No caso particular de  $F$  ser absolutamente contínua, von Mises encontrou as condições de suficiência que permitem simplificar o processo de averiguação se  $F$  pertence ou não, a um determinado domínio de atracção.

**Teorema 1.2.17** (Condições de von Mises)

Seja  $F$  uma f.d. absolutamente contínua,  $f$  a sua derivada e  $x_f$  o limite superior do seu suporte, então,

i.  $F \in D(\Lambda)$  se

a.  $x_f \leq +\infty$ ,

b.  $\exists x_0: f'(x) < 0, \forall x \in [x_0, x_f)$ ,

c.  $\lim_{t \rightarrow x_f} \frac{f'(t)(1-F(t))}{f^2(t)} = -1$ ;

ii.  $F \in D(\Phi_\alpha)$  se

a.  $x_f = +\infty$ ,

b.  $\lim_{t \rightarrow \infty} \frac{tf(t)}{1-F(t)} = \alpha > 0$ ;

iii.  $F \in D(\Psi_\alpha)$  se

a.  $x_f < +\infty$ ,

b.  $\lim_{t \rightarrow x_f} \frac{(x_f - t)f(t)}{1-F(t)} = \alpha > 0$ .

*Dem. in Resnick (1987), p. 63-66.*

Os dois próximos resultados estão essencialmente relacionados com distribuições discretas.



**Teorema 1.2.18** (Condição Necessária)

Se existir algum par  $c_n > 0$ ,  $d_n \in \mathfrak{R}$  para a d.f.  $F(x)$ , tal que

$$F^n(c_n x + d_n) \xrightarrow{n \rightarrow \infty} H(x)$$

onde  $H(x)$  assume uma das formas do Teorema 1.1.1, então,

$$\lim_{x \rightarrow x_f} \frac{F(x+0) - F(x)}{1 - F(x)} = 0.$$

Dem. in Galambos (1987), p. 84-85.

**Corolário 1.2.19**

Seja  $X$  uma v.a. discreta,  $X \in \mathfrak{N}_0^+$  e  $\text{Prob}(X = k) = p_k$ . Se  $\lim_{k \uparrow +\infty} p_k / \sum_{j=k}^{+\infty} p_j$  tender para zero, então não existem  $c_n > 0$ ,  $d_n \in \mathfrak{R}$ , tal que  $c_n^{-1}(M_n - d_n)$  convirja uma distribuição não degenerada.

As três distribuições estáveis para extremos podem ser representadas em uma única forma paramétrica.

**Definição 1.2.20** (forma de Jenkinson-von Mises ou função generalizada dos valores extremos)

Define-se  $H_\xi$  por

$$H_\xi(x) = \begin{cases} \exp\{-(1 + \xi x)^{-1/\xi}\}, & 1 + \xi x > 0 & \xi \neq 0 \\ \exp\{-\exp\{-x\}\}, & x \in \mathfrak{R} & \xi = 0 \end{cases}.$$

Existem outros resultados importantes em Teoria de Valores Extremos que se escreve de seguida.

**Teorema 1.2.21** (Caracterização de  $D(H_\xi)$ )

Para  $\xi \in \mathfrak{R}$  as seguintes condições são equivalentes

- i.  $F \in D(H_\xi)$
- ii. Existe uma função  $a(\cdot)$  positiva e mensurável tal que para  $1 + \xi x > 0$ ,

$$\lim_{x \rightarrow x_f} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = \begin{cases} (1 + \xi x)^{-1/\xi} & , \xi \neq 0 \\ e^{-x} & , \xi = 0 \end{cases} \quad (1.2)$$

- iii. Para  $x, y > 0$ ,  $y \neq 1$ ,

$$\lim_{s \rightarrow \infty} \frac{U(sx) - U(s)}{U(sy) - U(s)} = \begin{cases} \frac{x^\xi - 1}{y^\xi - 1} & , \xi \neq 0 \\ \frac{\ln x}{\ln y} & , \xi = 0 \end{cases}$$

Ao ii. pode ser atribuído uma interpretação estatística rescrevendo

$$\lim_{u \rightarrow x_f} \text{Prob} \left( \frac{X - u}{a(u)} > x \mid X > u \right) = \begin{cases} (1 + \xi x)^{-1/\xi} & , \xi \neq 0 \\ e^{-x} & , \xi = 0 \end{cases} \quad (1.3)$$

que é a distribuição assintótica dos excessos em relação a um determinado nível elevado,  $u$ , com escala  $a(u)$ .

Em (1.3) utiliza-se o conceito de excesso que se formaliza do seguinte modo.

**Definição 1.2.22** (Função de Distribuição Excesso e Função Excesso Médio)

Seja  $X$  uma v.a. com f.d.  $F$  com o lado direito do suporte  $x_f$ . Para um nível  $u < x_f$  fixo

$$F_u(x) = \text{Prob}\{X - u \leq x | X > u\}, \quad x \geq 0,$$

é a f.d. excesso de  $X$  sobre o nível  $u$ . A função

$$e(u) = E(X - u | X > u)$$

é a função excesso médio de  $X$ .

Na actividade seguradora  $F_u(x)$  pode ser interpretado por distribuição *excess-of-loss*.

### 1.3. Distribuição de Pareto Generalizada

A seguinte definição é motivada pela importância do resultado da expressão (1.3).

**Definição 1.3.1** (Distribuição de Pareto Generalizada)

Define-se  $G_\xi$  por

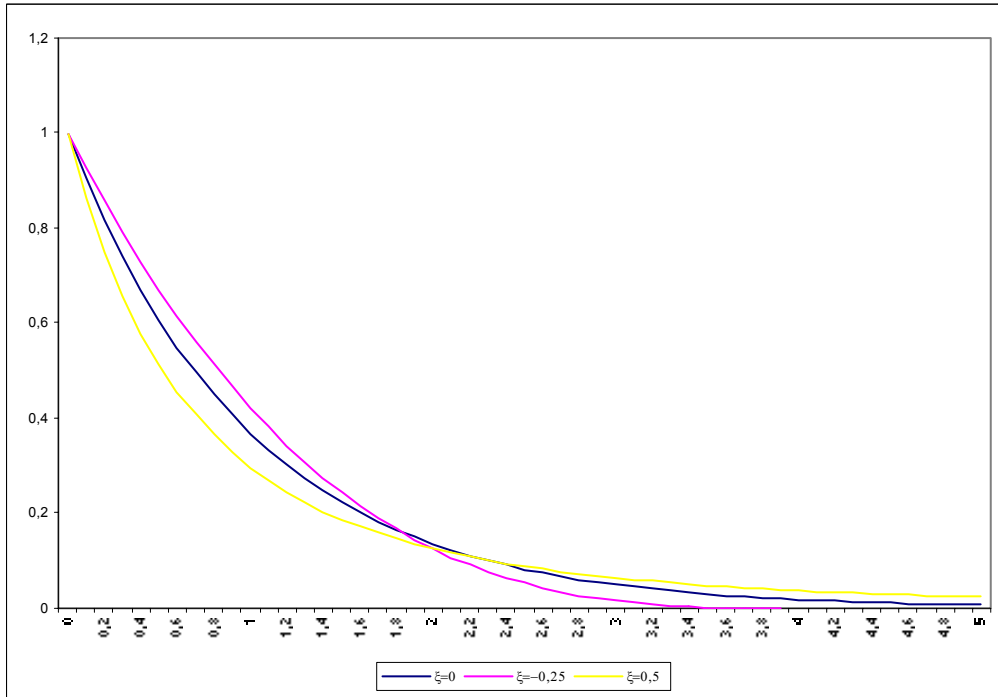
$$G_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi}, & \xi \neq 0 \\ 1 - e^{-x}, & \xi = 0 \end{cases}$$

onde

$$\begin{aligned} x &\geq 0, & \text{se } \xi \geq 0 \\ 0 \leq x &\leq -1/\xi, & \text{se } \xi < 0 \end{aligned}$$

É possível introduzir os parâmetros de localização e de escala,  $G_{\xi;\nu,\beta}$ , substituindo o  $x$  por  $(x - \nu)/\beta$ ,  $\nu \in \mathbb{R}$  e  $\beta > 0$ . No caso de  $\nu = 0$ ,  $G_{\xi;0,\beta}$  é substituído por  $G_{\xi;\beta}$ .

No gráfico que se segue estão representadas as funções de densidade para diferentes valores de  $\xi$ .



**Figura 1.3.1** - Funções de densidade das formas possíveis de  $G(\cdot)$ , para diferentes valores de  $\xi$ , com  $\beta = 1$  e  $\nu = 0$ .

A Distribuição de Pareto Generalizada (DPG) possui algumas propriedades de enorme interesse.

**Teorema 1.3.2** (Propriedades da DPG)

- i. Se a f.d. de  $X$  for DPG com parâmetros  $\xi$  e  $\beta$ , então,  $E(X) < \infty$  se e só se  $\xi < 1$ ;
- ii. Para qualquer  $\xi \in \mathbb{R}$ ,  $F \in D(H_\xi)$  se e só se

$$\lim_{u \rightarrow x_f} \sup_{0 < x < x_f - u} |F_u(x) - G_{\xi, \beta(u)}(x)| = 0,$$

ou seja,  $\bar{F}_u(x) = \text{Prob}(X - u > x | X > u) \approx \bar{G}_{\xi, \beta(u)}(x)$ ,  $u \rightarrow x_f$ ,  $x > 0$ ;

iii. Se  $x_1, x_2 \in D(\xi, \beta)$ ,  $i = 1, 2$ , então

$$\frac{\overline{G}_{\xi; \beta}(x_1 + x_2)}{\overline{G}_{\xi; \beta}(x_1)} = \overline{G}_{\xi; \beta + \xi x_1}(x_2)$$

ou seja, a probabilidade de  $X$  exceder  $x_1 + x_2$  uma vez que excede  $x_1$ , continua a ser uma DPG. Esta propriedade implica que a DPG é fechada, trata-se de uma propriedade de enorme interesse nos tratados excess-of-loss e stop-loss.

iv. Seja  $N \sim \text{Poi}(\lambda)$  independente da sucessão i.i.d.  $X_n \sim G_{\xi; \beta}$ , seja ainda

$$M_n = \max(X_1, X_2, \dots, X_N), \text{ então}$$

$$\text{Prob}(M_N \leq x) = \exp \left\{ -\lambda \left( 1 + \xi \frac{x}{\beta} \right)^{-1/\xi} \right\} = H_{\xi; \mu, \psi}(x)$$

onde  $\mu = \beta \xi^{-1} (\lambda^\xi - 1)$  e  $\psi = \beta \lambda^\xi$ .

Por outras palavras, nas condições referidas  $\text{Prob}(M_N \leq x)$  tem como domínio de atracção GEV.

v.  $X_n \sim G_{\xi; \beta}$ ,  $\xi < 1$ , então para  $u < x_f$ ,

$$e(u) = E(X - u | X > u) = \frac{\sigma + \xi u}{1 - \xi}, \quad \beta + u\xi > 0,$$

por conseguinte, a função de excesso médio é linear. Trata-se de uma propriedade com um enorme sentido prático.

Um dos problemas na modelização de DPG é a escolha do nível de  $u$ . Alguns autores propõem a escolha de um nível de  $u$  a partir do qual, a função de excesso médio empírica se comporta como uma função linear, é um tema a ser aprofundado no Capítulo 3 deste trabalho.



## CAPÍTULO 2

---

### Análise e Apresentação dos Dados

Pretende-se, no presente capítulo apresentar e analisar algumas técnicas para conhecimento do comportamento de dados de cauda pesada. É dada maior atenção às indemnizações de elevado montante, sendo introduzidos alguns conceitos relevantes.

Os dados utilizados são retirados da carteira de seguros de danos patrimoniais de empresas de uma filial da Allianz Group. O período de observação foi de 01-01-1993 a 31-03-2001 e todos os valores foram corrigidos para preços de 2002.

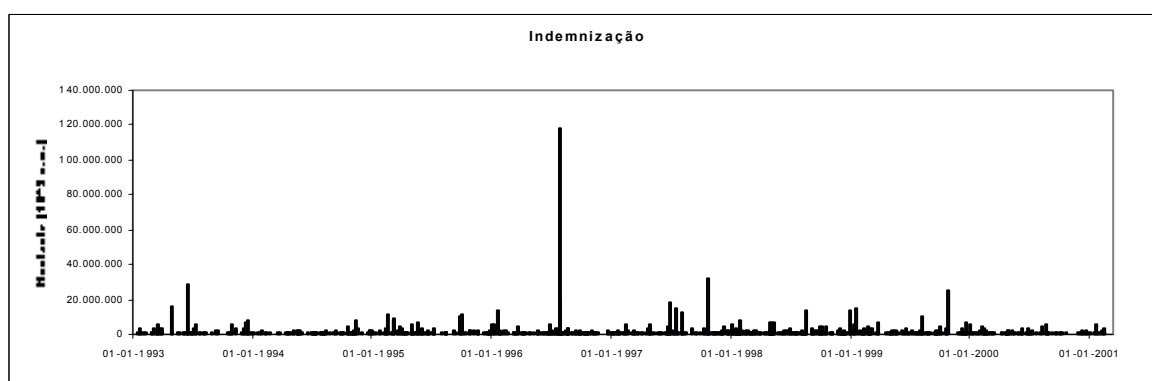
Apresentam-se de seguida algumas estatísticas dos dados em análise.

Média	645.090
Desvio Padrão	2.514.059
Mínimo	130.019
1º Percentil	130.877
1º Quartil	178.033
Mediana	271.391
3º Quartil	532.543

99º Percentil	5.855.928
Máximo	117.953.980
N	3296

Através do quadro anterior é possível constatar a existência de um pequeno conjunto de indemnizações de elevado montante.

Na Figura 2.1, verifica-se que as indemnizações de maior montante estão distribuídas regularmente ao longo do período, sem que apresentem alguma tendência na evolução das indemnizações. Assim, parece razoável assumir a hipótese i.i.d..



**Figura 2.1** – *Comportamento dos sinistros extremos verificados entre 1993 e 2001.*

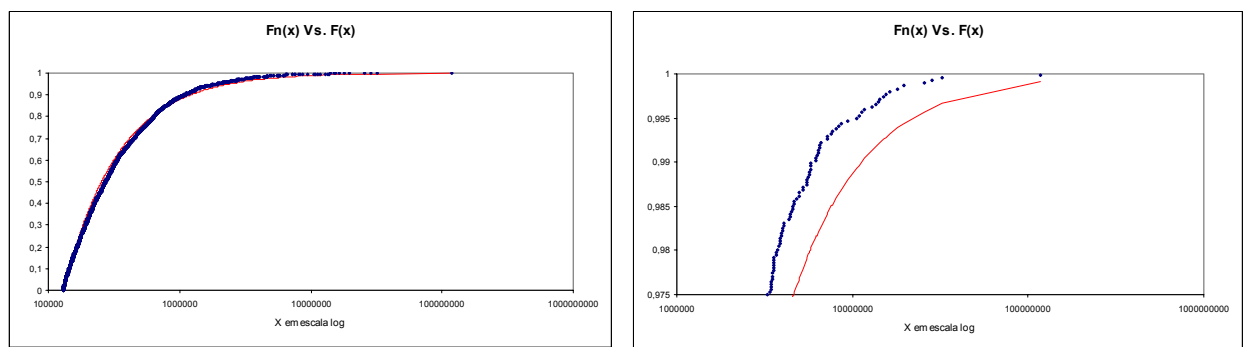
Numa primeira análise, os dados foram ajustados a uma distribuição de Pareto, na forma de  $F(x) = 1 - (\beta/x)^\alpha$ ,  $x > \beta$ ,  $\beta = 130000$ , que é igual ao valor da franquía, e  $\alpha = 1,03233$ , estimado pelo método de máxima verosimilhança.

Pela Figura 2.2, constata-se que a distribuição de Pareto tende a sobrestimar a cauda, ainda que o ajustamento pareça relativamente bom para os valores intermédios. Esse

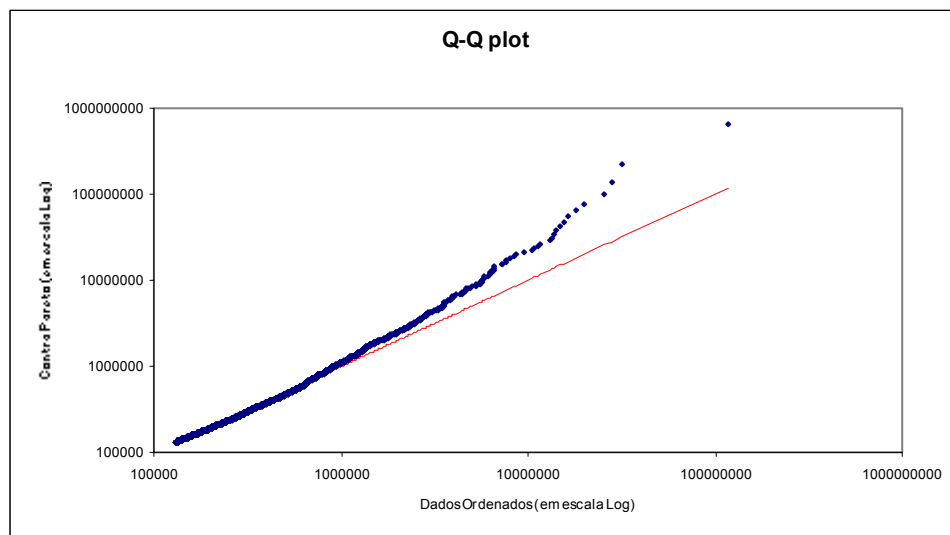


efeito torna-se mais evidente na Figura 2.3, que permite visualizar a diferença dos quantis estimados.

A sobrestimação da cauda pode traduzir-se em custos financeiros, por exemplo, na determinação do prémio por *layer* num tratado de *excess-of-loss* a pagar à resseguradora ou em custos económicos na determinação do nível adequado do capital da seguradora.



**Figura 2.2** – No gráfico da esquerda estão representadas a distribuição empírica e a distribuição de Pareto. O gráfico da direita tem a escala ampliada na cauda.



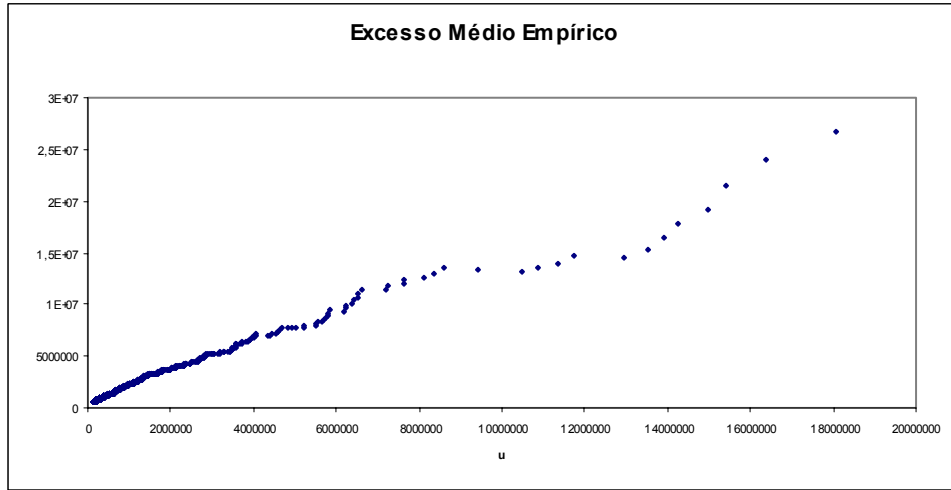
**Figura 2.3** – Comparação com os quantis da distribuição de Pareto.

O cálculo do excesso médio empírico fornece algumas ideias sobre o comportamento da cauda. Na Figura 2.4 está representado

$$\{(u, e_n(u)), X_{n,n} < u < X_{1,n}\}$$

onde

$$e_n(u) = \frac{\sum_{i=1}^n (X_i - u)^+}{\sum_{i=1}^n I_{\{X_i > u\}}}$$



**Figura 2.4** - Função de excesso médio empírico.

Para os dados em análise, a função de excesso médio empírico é crescente com  $u$ . Esta constatação confirma que a função de distribuição empírica tem uma cauda pesada. Por outro lado, verifica-se que a função excesso médio assume uma forma linear sensivelmente a partir de  $u = 5500000$  ou de  $u = 6000000$ , indiciando que, a partir desses níveis, a cauda pode ser modelada com uma Distribuição de Pareto Generalizada, como sugerido no Capítulo 2.

Outra questão intimamente ligada ao comportamento da cauda, é a existência de momentos. Uma forma de averiguar a sua existência consiste em calcular

$$R_n(p) = \frac{M_n(p)}{S_n(p)}, \quad n \geq 1, \quad p \geq 0,$$

onde,  $S_n(p) = |X_1|^p + \dots + |X_n|^p$  e  $M_n(p) = \max(|X_1|^p, \dots, |X_n|^p)$ .

Se

$$R_n(p) \xrightarrow{q.c.} 0 \quad \Leftrightarrow \quad E(|X|^p) < \infty$$

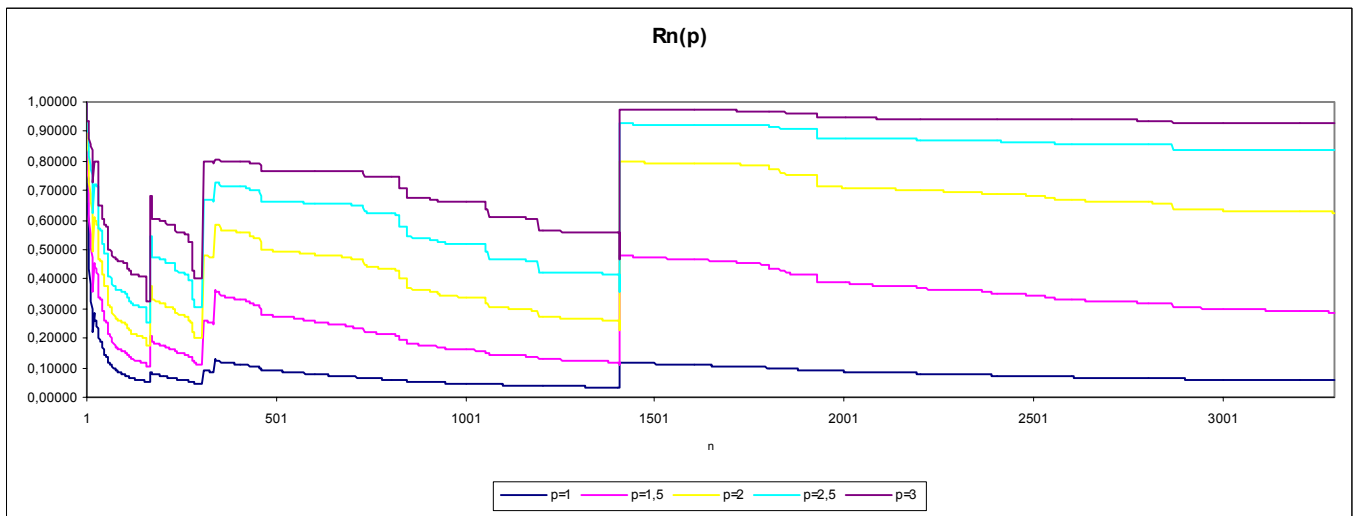
$$R_n(p) \xrightarrow{P} 0 \quad \Leftrightarrow \quad E(|X|^p I_{\{|X| < x\}}) \in R_0$$

$$R_n(p) \xrightarrow{d} Y_2(p) \quad \Leftrightarrow \quad P(|X| > x) \in R_{-\alpha}, \quad 0 < \alpha < 1$$

$$R_n(p) \xrightarrow{P} 1 \quad \Leftrightarrow \quad P(|X| > x) \in R_0$$

onde,  $Y_2(p)$  é uma v.a. apropriada e não degenerada. Informação detalhada pode ser encontrada em Embrechts *et al.* (1997).

Na Figura 2.5 estão representados os  $R_n(p)$  para vários valores de  $p$ . Verifica-se que, para  $p > 1,5$ ,  $\lim_{n \uparrow \infty} R_n(p)$  não tende para zero, indiciando a inexistência dos momentos de ordem superior.



**Figura 2.5 -  $R_n(p)$  para vários valores de  $p$ .**



## CAPÍTULO 3

---

### Modelação dos dados

Pretende-se, ao longo deste capítulo, modelar os dados apresentados no Capítulo 3 de acordo com as metodologias relacionadas com a Teoria de Valores Extremos, a fim de estudar com maior profundidade, o comportamento das indemnizações extremas. Assim, na Secção 3.1, os dados são modelados por ajustamento de uma Distribuição de Pareto Generalizada, seguindo a metodologia apresentada em Embrechts *et al.* (1997) e McNeil (1997). Na Secção 3.2.1 é apresentada uma abordagem semi-paramétrica, sendo introduzido um método adaptativo, seguindo as metodologias apresentadas em Danielsson *et al.* (2001) e Gomes *et al.* (2002) na Secção 3.2.2. No final, na Secção 3.3, procede-se ao desenvolvimento da discussão sobre ambas as abordagens.

#### 3.1 Modelação pela Distribuição de Pareto Generalizada

Pretende-se nesta Secção estudar o comportamento da cauda seguindo uma abordagem paramétrica. Na abordagem paramétrica procura-se ajustar uma distribuição aos excessos,  $Y = X - u$ ,  $X > u$ , através da aplicação do resultado *ii.* do Teorema 1.3.2, onde se afirma: à medida que  $u \rightarrow x_f$ ,  $\bar{F}_u(x) = \text{Prob}(X - u > x | X > u) \approx \bar{G}_{\xi; \beta(u)}(x)$ .

A determinação do nível óptimo de  $u$  é essencial para este tipo de modelação. Ao contrário do que acontece na abordagem semi-paramétrica, que é apresentada na Secção 3.2, ainda não existe nenhum método adaptativo para a escolha do nível de óptimo de  $u$ , aceite de uma forma generalizada.

Na determinação do nível de  $u$  é necessário ter em conta o seguinte dilema:

- i. uma escolha do nível de  $u$  demasiado elevado, pode conduzir a uma maior variância nas estimativas, na medida em que o número de observações que excedem  $u$  é reduzido;
- ii. ao passo que, uma escolha do nível de  $u$  demasiado baixo, pode originar a um maior viés, além de não se poder aplicar o *ii.* do Teorema 2.3.2.

No Capítulo 2 foi sugerido uma forma de determinar o nível de  $u$ , que consiste na escolha de um valor de  $X$ , à direita do qual, a função de excesso médio empírica se assemelha a uma função linear. Existem, no entanto, situações em que o nível óptimo não é evidente, sendo vários os níveis aceitáveis, como, por exemplo, os indicados na Figura 2.4 para os dados em análise.

Na Figura 2.4, atendendo ao número de observações e ao valor do nível em si, existem alguns indícios de que o nível óptimo possa ser  $u = 550000$  ou  $u = 600000$ .

Os parâmetros  $\xi$  e  $\beta(u)$  podem ser estimados pelo método de máxima verosimilhança. Seja  $F$  uma Distribuição de Pareto Generalizada com parâmetros  $\xi$  e  $\beta > 0$ , a função de densidade  $f$  é

$$f(x) = \frac{1}{\beta} \left( 1 + \xi \frac{x}{\beta} \right)^{-\frac{1}{\xi}-1}, \quad x \in D(\xi, \beta).$$

onde

$$D(\xi, \beta) = \begin{cases} [0, \infty), & \xi \geq 0 \\ [0, -\beta/\xi], & \xi < 0 \end{cases}$$

A função log-verosimilhança é

$$\ell((\xi, \beta); \underline{X}) = -n \ln \beta - \left( \frac{1}{\xi} + 1 \right) \sum_{i=1}^n \ln \left( 1 + \frac{\xi}{\beta} X_i \right).$$

Para  $\xi > -1/2$ , é possível demonstrar que

$$n^{1/2} \left( \hat{\xi}_n - \xi, \frac{\hat{\beta}_n}{\beta} - 1 \right) \xrightarrow{d} N(0, M^{-1}), \quad n \rightarrow \infty,$$

onde,

$$M^{-1} = (1 + \xi) \begin{pmatrix} 1 + \xi & -1 \\ -1 & 2 \end{pmatrix},$$

e  $\hat{\xi}_n$  e  $\hat{\beta}_n$  são estimadores de máxima verosimilhança.

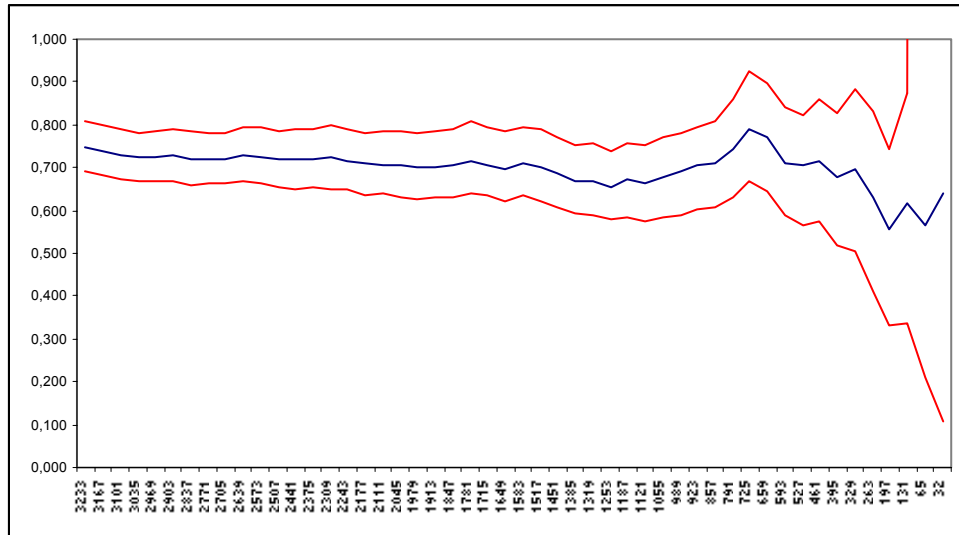
Por se pretender modelar a distribuição dos excessos de um determinado nível  $u$ , o vector  $\underline{X}$  é substituído pelo vector  $\underline{Y}$ , o vector dos excessos, e  $N$  por  $N_u$ , o número de observações que excederam  $u$ .

O estimador de  $\xi$  pode ser obtido após uma reparametrização.  $(\xi, \beta) \rightarrow (\xi, \tau)$ , onde  $\tau = -\xi / \beta$ . Resultando,

$$\hat{\xi} = \hat{\xi}(\tau) = N_u^{-1} \sum_{i=1}^{N_u} \ln(1 - \tau Y_i),$$

$$\tau \in (-\infty, \max(Y_1, \dots, Y_{N_u})): h(\tau) = \frac{1}{\tau} + \frac{1}{N_u} \left( \frac{1}{\hat{\xi}(\tau)} + 1 \right) \sum_{i=1}^{N_u} \frac{Y_i}{1 - \tau Y_i} = 0.$$

Na Figura 3.1.1. estão representados os valores de  $\hat{\xi}$  para diferentes níveis de  $u$ . Estão representados também os intervalos de confiança *bootstrap* com 95% de confiança, com 250 réplicas. A partir da Figura 3.1.1, é possível constatar que o valor de  $\hat{\xi}$  é relativamente estável dentro do intervalo (0,6;0,8).



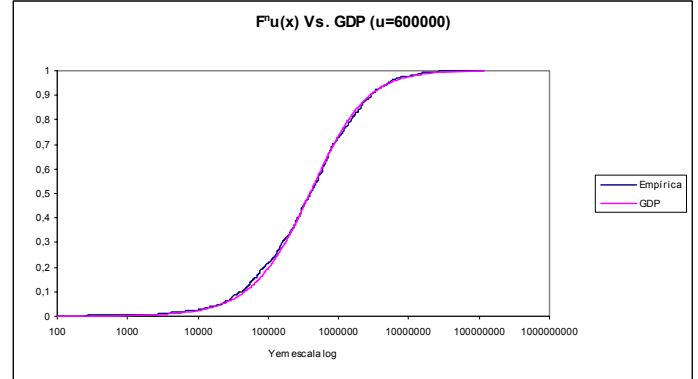
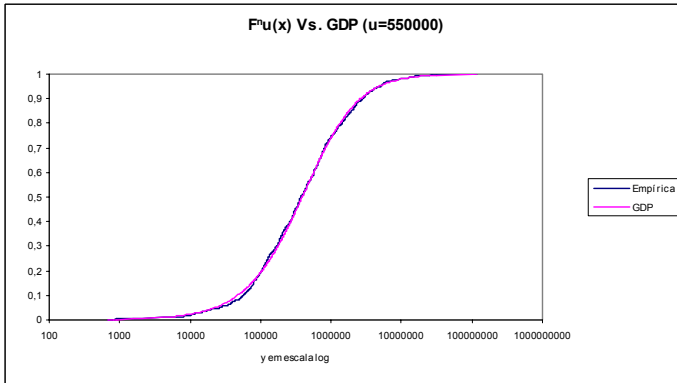
**Figura 3.1.1** – Valores de  $\hat{\xi}$ , da DPG, para diferentes níveis de  $u$ .

O valor das estimativas estão apresentados no quadro seguinte:

$u$	550.000	600.000
$\hat{\xi}$	0,74139	0,79825
$\hat{\beta}$	430.802	422.063
$n_u$	898	732

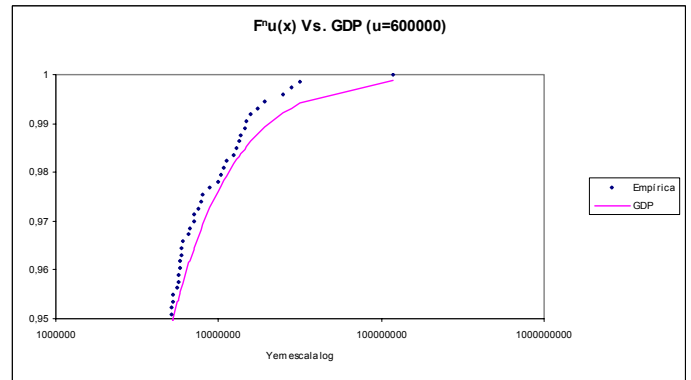
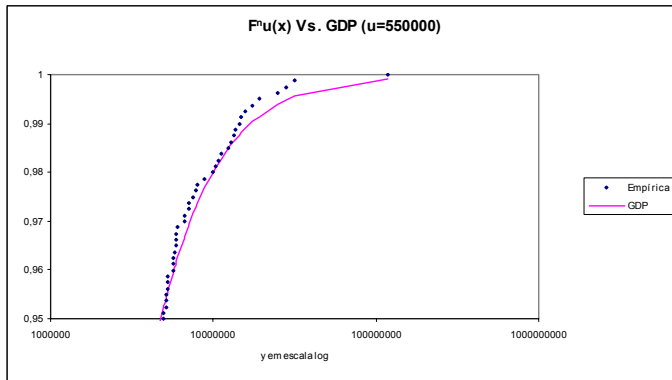
Nos gráficos seguintes estão representadas as distribuições condicionadas e as respectivas funções de distribuição empírica condicionada, para  $u = 550000$  e  $u = 600000$ .





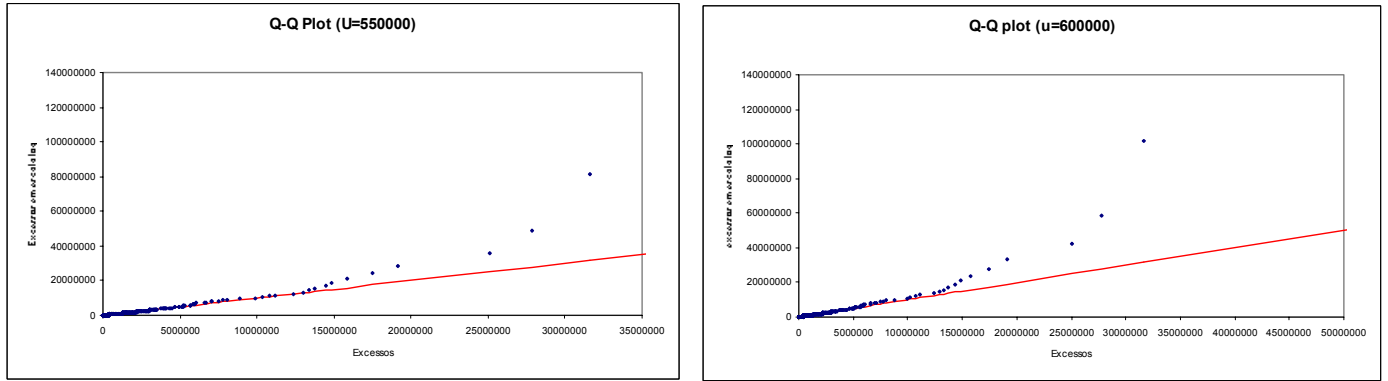
**Figura 3.1.2** – Função de distribuição ajustada aos excessos com  $u = 550.000$  (à esquerda) e  $u = 600.000$  (à direita).

Os gráficos que se seguem têm a escala ampliada no sentido de dar maior enfoque à cauda.



**Figura 3.1.3** – Função de distribuição ajustada aos excessos com  $u = 550.000$  (à esquerda) e  $u = 600.000$  (à direita), com a cauda ampliada.

Em ambos os modelos o ajustamento parece ser razoável. A mesma conclusão pode ser retirada pela observação dos gráficos Q-Q.



**Figura 3.1.4** – Comparação dos quantis dos excessos com as distribuições ajustadas com  $u = 550.000$  (à esquerda) e  $u = 600.000$  (à direita).

Após o ajustamento das funções de distribuição aos excessos, através da distribuição de Pareto Generalizada, é necessário reparametrizar os modelos por forma a obter a  $F(x)$ , ou seja, as distribuições não condicionadas. Assim,

$$\hat{F}(x) = 1 - \left( 1 + \hat{\xi} \frac{x - u - \hat{v}}{\hat{\beta}'} \right)^{-1/\hat{\xi}}, \quad x \geq u,$$

onde

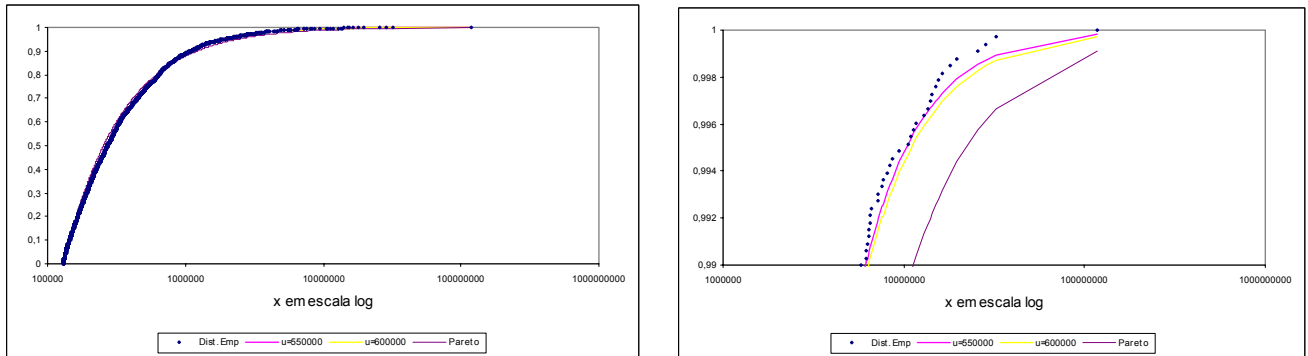
$$\hat{v} = \frac{\hat{\beta}}{\hat{\xi}} \left( \left( \frac{N_u}{n} \right)^{\hat{\xi}} - 1 \right)$$

e

$$\hat{\beta}' = \hat{\beta} \left( \frac{N_u}{n} \right)^{\hat{\xi}}.$$

Para  $x < u$ , o número de observações é relativamente elevado: 2398 para  $u = 550000$  e 2564 para  $u = 600000$  e a diferença entre duas estatísticas de ordem consecutivas é pequena, o que leva a considerar que  $X$  pode ser modelado com a função de distribuição empírica.

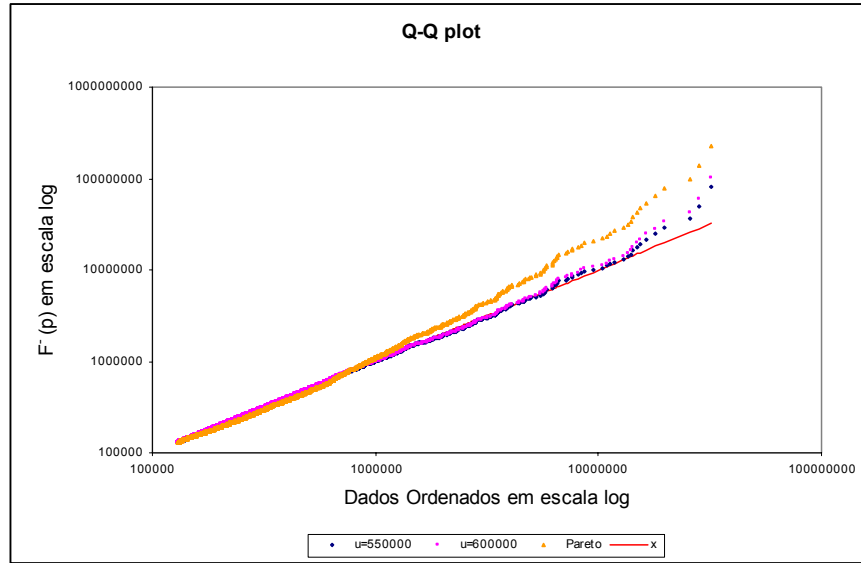
Nos gráficos seguintes são representadas as várias  $F(x)$  para os dois valores de  $u$ . É igualmente representada a distribuição de Pareto apresentada no Capítulo 3, como termo de comparação.



**Figura 3.1.5** – No gráfico da esquerda estão representadas a distribuição empírica e a distribuição ajustada com  $u = 550.000$  e  $u = 600.000$ . O gráfico da direita tem a escala ampliada na cauda.

O gráfico da esquerda sugere que o ajustamento de todos os modelos é aceitável. No gráfico da direita onde a escala é ampliada, verificam-se melhorias no ajustamento relativamente à distribuição de Pareto. As mesmas conclusões podem ser retiradas da Figura 3.1.6, onde estão representados os gráficos Q-Q.

Existe, em ambos os modelos, uma sobrestimação dos quantis em relação à distribuição empírica. Trata-se de um resultado esperado dado que um dos objectivos deste tipo de análise, é a estimação dos quantis fora da amostra.



**Figura 3.1.6** – Comparação dos quantis com os da distribuição ajustada com  $u = 550.000$  e  $u = 600.000$ .

### 3.2 Abordagem semi paramétrica

Nesta Secção, a modelação da cauda da distribuição, segue uma abordagem semi-paramétrica. Em 3.2.1 é apresentada a ideia subjacente a esta abordagem. Na Secção 3.2.2 é apresentado um algoritmo adaptativo para estimar a fracção óptima da amostra a utilizar.

#### 3.2.1 Estimação Sob Condição de Domínio de Atracção para Máximos

Seja  $X_1, X_2, \dots, X_n$  uma sequência de v.a. i.i.d. de uma f.d.  $F \in D(\Phi_\alpha)$ , tendo em conta o Teorema 1.2.7, a cauda da distribuição pode ser escrita da seguinte forma:

$$\bar{F}(x) = x^{-\alpha} L(x), \quad x > 0,$$

onde  $L$  é uma função de variação lenta.

Por outro lado, atendendo ao Teorema 1.2.2, se  $F \in D(H_\xi)$  então

$$\lim_{n \rightarrow \infty} n\bar{F}(c_n x + d_n) = -\ln H_\xi(x).$$

Seja  $u = c_n x + d_n$ , a expressão em cima pode ser rescrita do seguinte modo

$$n\bar{F}(u) \approx \left(1 + \xi \frac{u - d_n}{c_n}\right)^{-1/\xi}.$$

Desta forma o estimador da cauda é

$$\hat{\bar{F}}(u) = \frac{1}{n} \left(1 + \hat{\xi} \frac{u - \hat{d}_n}{\hat{c}_n}\right)^{-1/\hat{\xi}}$$

com estimadores adequados de  $\hat{\xi}$ ,  $\hat{c}_n$  e  $\hat{d}_n$ .

A estimação de  $\xi$  é baseada nas  $k$  maiores estatísticas de ordem, e é desenvolvida na Secção 3.2.2 com maior profundidade. A dimensão de  $k$  deve satisfazer duas condições:  $k(n) \rightarrow \infty$  e  $n/k(n) \rightarrow \infty$ , que pode ser interpretada da seguinte forma: o número de estatísticas de ordem deve ser suficientemente elevado, contudo, há que ter em conta que apenas a cauda é analisada. Uma escolha adequada do valor de  $k$  permite obter algumas propriedades como a consistência e a normalidade assintótica do estimador.

Em relação aos estimadores de  $\hat{c}_n$  e  $\hat{d}_n$ , estes sofrem uma ligeira alteração relativamente ao apresentado no Teorema 1.2.7, onde se definiu  $c_n = F^{\leftarrow}(1 - n^{-1})$  que no presente modelo é substituído por  $c_{n/k} = F^{\leftarrow}(1 - (n/k)^{-1})$ , onde  $n/k$  satisfaz as condições enunciadas.

Embora existam vários estimadores de  $\alpha$  propostos na literatura de Teoria de Valores Extremos, optou-se no presente trabalho por utilizar apenas o estimador de Hill, por ser, por um lado, um estimador natural e, por outro, pela sua importância histórica.

Seja  $X$  uma v.a. com a f.d.  $F$  tal que,

$$\text{Prob}(X > x) = \bar{F}(x) = x^{-\alpha}, \quad x \geq 1, \quad \alpha > 0.$$

Então,  $Y = \ln X$  tem como distribuição

$$\text{Prob}(Y > y) = e^{-\alpha y}, \quad y \geq 0, \quad \alpha > 0,$$

ou seja,  $Y \sim \text{Exp}(\alpha)$ . Assim, o estimador de máxima verosimilhança é:

$$\hat{\alpha} = \left( \frac{1}{n} \sum_{j=1}^n \ln X_j \right)^{-1} = \left( \frac{1}{n} \sum_{j=1}^n \ln X_{j,n} \right)^{-1}.$$

Generalizando o resultado apresentado

$$\bar{F}(x) = Cx^{-\alpha}, \quad x \geq u > 0$$

onde  $C = u^\alpha$ , para um nível  $u$  conhecido, o estimador assumirá a seguinte forma:

$$\hat{\alpha}_n = \left[ \frac{1}{n} \sum_{j=1}^n \ln \left( \frac{X_{j,n}}{u} \right) \right]^{-1} = \left( \frac{1}{n} \sum_{j=1}^n \ln X_{j,n} - \ln u \right)^{-1}$$

Contudo, na prática, o nível exacto de  $u$  não é conhecido. Mas se  $F \in D(\Phi_\alpha)$ , então  $\bar{F}$  tem um comportamento que se assemelha à cauda da distribuição de Pareto.

Seja

$$K = \# \{i : X_{i,n} > u, i = 1, \dots, n\}$$

Reparametrizando os estimadores, substituindo o suposto nível  $u$  pela estatística de ordem  $\{K = k\}$ , o estimador pode ser rescrito como a seguir se representa:

$$\hat{\alpha}^{(H)} = \hat{\alpha}_{k,n}^{(H)} = \left( \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1} \quad (3.1)$$

e

$$\hat{C}_{k,n} = \frac{k}{n} X_{k,n}^{\hat{\alpha}_{k,n}^{(H)}}$$

onde  $k = k(n)$  satisfaz as condições referidas. A expressão (3.1) é conhecida como estimador de Hill.

O estimador da cauda é

$$\hat{\bar{F}}(x) = \frac{k}{n} \left( \frac{x}{X_{k,n}} \right)^{-\hat{\alpha}_{k,n}^{(H)}},$$

isto é,  $\bar{F}$  tem um comportamento de tipo Pareto acima do nível aleatório  $X_{k,n}$ .

O próximo teorema apresenta as propriedades do estimador de Hill.

**Teorema 3.2.1** (Propriedades do Estimador de Hill)

Seja  $\{X_n\}$  estritamente estacionária com distribuição marginal  $F$  satisfazendo para todo o  $\alpha > 0$  e  $L \in R_0$ ,

$$\bar{F}(x) = x^{-\alpha} L(x), \quad x > 0,$$

Seja  $\hat{\alpha}^{(H)} = \hat{\alpha}_{k,n}^{(H)}$  o estimador de Hill.

- i. (consistência fraca) se  $\{X_n\}$  é i.i.d. e se  $k(n) \rightarrow \infty$  e  $k(n)/n \rightarrow 0$  para  $n \rightarrow \infty$ , então,

$$\hat{\alpha}^{(H)} \xrightarrow{p} \alpha$$

- ii. (consistência forte) se  $\{X_n\}$  é i.i.d. e se  $k(n)/n \rightarrow 0$ ,  $k(n)/\ln \ln n \rightarrow \infty$  para  $n \rightarrow \infty$ , então,

$$\hat{\alpha}^{(H)} \xrightarrow{q.c.} \alpha$$

- iii. (normalidade assintótica) Se  $\{X_n\}$  é i.i.d. e  $k(n) \rightarrow \infty$  a uma velocidade apropriada, então,

$$\sqrt{k}(\hat{\alpha}^{(H)} - \alpha) \xrightarrow{d} N(0, \alpha^2).$$

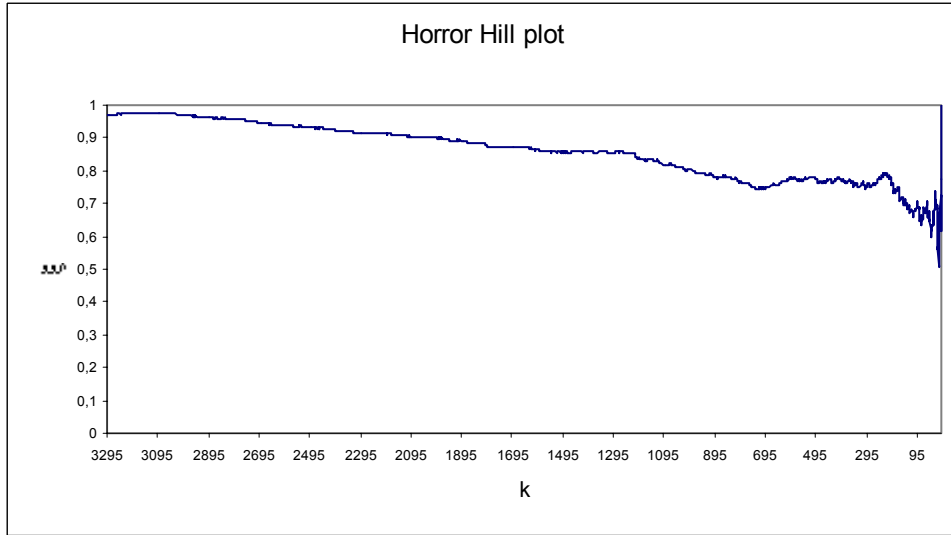
### 3.2.2 Determinação do Nível Ótimo de $k$

A determinação do nível ótimo de  $k$  é um dos temas de maior interesse em Teoria de Valores Extremos. A escolha do valor de  $k$  condiciona o valor das estimativas e as propriedades do estimador. À semelhança do que acontece na determinação do valor do nível  $u$ , na modelação pela Distribuição de Pareto Generalizada, ao determinar o valor de  $k$  o analista enfrenta o seguinte dilema:

- i. para garantir que apenas a cauda seja analisada deve-se escolher um valor de  $k$  relativamente baixo. Contudo, a variância das estimativas pode aumentar se se seleccionar um  $k$  excessivamente baixo.
- ii. porém, ao aumentar demasiado o valor de  $k$ , poderá aumentar o viés.

No gráfico da Figura 3.2.1 está representado o *horror Hill plot*, através do qual é possível visualizar a sensibilidade do valor de  $\hat{\xi} = 1/\hat{\alpha}$ , em função dos valores de  $k$ .





**Figura 3.2.1** – *Horror Hill plot*

Embrechts *et al.* (1997) propõe o valor de  $k$  dentro do intervalo onde, para diferentes valores de  $k$ , o valor de  $\hat{\xi}$  se mantém relativamente estável. Trata-se de um critério relativamente vago. Todavia, existem alguns autores que recorrem a métodos adaptativos.

A abordagem que se apresenta é baseada nos trabalhos de Danielsson *et al.* (2001) e Gomes *et al.* (2002), que é desenvolvida no âmbito da condição de segunda ordem,

$$\lim_{t \rightarrow \infty} \frac{U(tx)/U(t) - x^\xi}{A(t)} = x^\xi \frac{x^\rho - 1}{\rho} \quad (3.2)$$

onde,  $A(t) \in RV_\rho$  e  $U(t) = F^\leftarrow(1 - 1/t)$  é uma função mensurável com sinal constante e  $\rho \leq 0$ , que é o parâmetro de segunda ordem que determina a convergência de  $\bar{F}(tx)/\bar{F}(x)$  para  $t^{-\alpha}$ . À medida que  $|\rho|$  aumenta, a convergência à aproximação de primeira ordem também aumenta.

Hall (1982) sugere que é possível obter  $k(n)$  através de

$$k_0(n) = \arg \min_k \text{AssimE}(\xi_n(k) - \xi)^2$$

i.e., escolher  $k_0(n)$  de modo a minimizar o erro quadrático médio assintótico (EQMA).

De onde resulta,

$$\sqrt{k_0(n)}(\xi_n(k_0(n)) - \xi) \xrightarrow{d} N(b, \xi^2)$$

ou seja, existe viés assintótico  $b$ .

Desta forma, é necessário determinar um estimador de  $\hat{k}_0(n)$  tal que,

$$\sqrt{\hat{k}_0(n)}(\xi_n(\hat{k}_0(n)) - \xi) \xrightarrow{d} N(b, \xi^2).$$

Para isso, Hall (1990) e Gomes (1994) sugerem o uso de metodologias de *bootstrap*, acrescentando uma hipótese adicional sobre  $A(t)$  em que

$$A(t) = ct^\rho, \quad \rho < 0. \quad (3.3)$$

No entanto, as metodologias de *bootstrap* habituais não garantem que o EQMA estimado convirja para o valor assintótico. Este problema é ultrapassado por uma reamostragem de dimensão  $n_1$ ,  $n_1 < n$ .

Assim, em vez de minimizar

$$\text{EQM}(n, k) = E\left[\left(\xi_n^*(k) - \xi_{n,aux}\right)^2 | \underline{X}_n\right],$$

onde  $\xi_{n,aux}$  é uma estimativa inicial que resulta de um estimador consistente e  $\xi_n^*(k)$  é

um estimador de *bootstrap* com a reamostragem de dimensão  $n$ , minimiza-se

$$\text{EQM}(n_1, k_1) = E\left[\left(\xi_{n_1}^*(k_1) - \xi_{n,aux}\right)^2 | \underline{X}_n\right].$$

**Teorema 3.2.2** (Sequência óptima de  $k_0^*(n_1)$ )

Se (3.2) e  $A(t) = ct^\rho$  verificarem com  $n_1 = O(n^{1-\varepsilon})$ ,  $0 < \varepsilon < 1$ , existe uma função  $D(\rho)$ ,

tal que a sequência de  $k_0^*(n_1)$  que minimiza  $EQM(n_1, k_1)$  é

$$k_0^*(n_1) = D(\rho)n_1^{2\rho/(1-2\rho)}(1 + o(1))$$

e a sequência de  $k_0(n)$  é

$$k_0(n) = D(\rho)n^{2\rho/(1-2\rho)}(1 + o(1)).$$

Como foi referido anteriormente, a metodologia *bootstrap* minimiza  $EQM(n_1, k_1)$ . O

que se obtém é  $k_0^*(n_1)$ , que é óptimo para uma amostra de dimensão  $n_1$ , mas o que se pretende estimar é  $k_0(n)$ . Pelo Teorema 3.2.2, a seguinte relação é verdadeira

$$\frac{k_0^*(n_1)}{k_0(n)} = \left(\frac{n_1}{n}\right)^{2\rho/(2\rho-1)}(1 + o(1)),$$

à medida que  $n \rightarrow \infty$ .

Porém,  $\rho$  não é conhecido e, para estimar  $k_0(n)$ , recorre-se a uma segunda amostra de dimensão  $n_2 = n_1^2/n$ . Novamente, pelas expressões do Teorema 3.2.2 obtém-se a seguinte relação

$$\frac{[k_0^*(n_1)]^2}{k_0^*(n_2)} = k_0(n)(1 + o(1)),$$

à medida que  $n \rightarrow \infty$ .

Assim,

$$\hat{k}_0(n; \xi_{n,aux}, n_1) = \frac{[\bar{k}_0^*(n_1)]^2}{\bar{k}_0^*(n_1^2/n)}$$

onde,  $\bar{k}_0^*(n_1)$  e  $\bar{k}_0^*(n_1^2/n)$  representam os valores óptimos obtidos pelas metodologias de *bootstrap*.

Os resultados que se apresentam seguidamente, seguiram o seguinte algoritmo para diferentes valores de  $n_1$ , expressos em percentagem de  $n$ :

1. Foi escolhido um valor inicial de  $\xi = 0,765$  que seria um valor aproximado se o valor de  $k$  fosse seleccionado segundo o critério de Embrechts et al. (1997).
2. Retiraram-se 1000 ( $B$ ) amostras *bootstrap*  $(x_1^*, \dots, x_{n_2}^*)$  e  $(x_1^*, \dots, x_{n_2}^*, x_{n_2+1}^*, \dots, x_{n_1}^*)$ ,
3. Para cada amostra foi calculado para cada  $k$  o valor da estatística

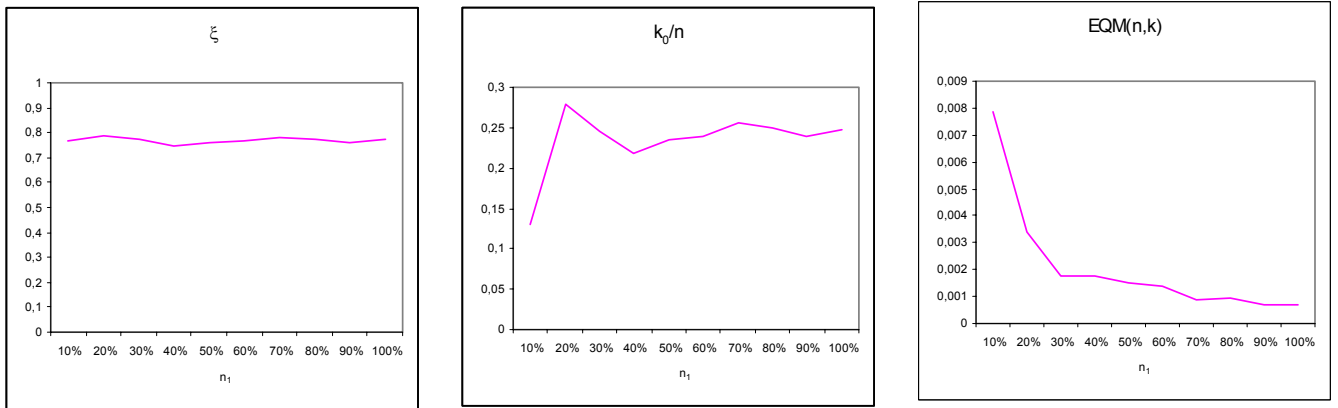
$$T_{n_i}(k) = \xi_{n_i}^*(k) - \xi_{n,aux}, \quad i = 1, 2$$

4. Obteve-se  $\bar{k}_0^*(n_i) = \arg \min_k EQM(n_i, k)$ ,  $i = 1, 2$ , onde

$$EQM(n_i, k) = \frac{1}{B} \sum_{l=1}^B t_{n_i, l}(k)^2$$

5. Foi calculado  $\hat{k}_0(n, \xi_{n,aux}, n_1) = \frac{[\bar{k}_0^*(n_1)]^2}{\bar{k}_0^*(n_2)}$
6. Por fim, foi obtido o valor de  $\xi_{n, n_1}^* = \xi_n(\hat{k}_0(n, \xi_{n,aux}, n_1))$ .

Os resultados obtidos estão sumarizados nos gráficos da Figura 3.2.2.

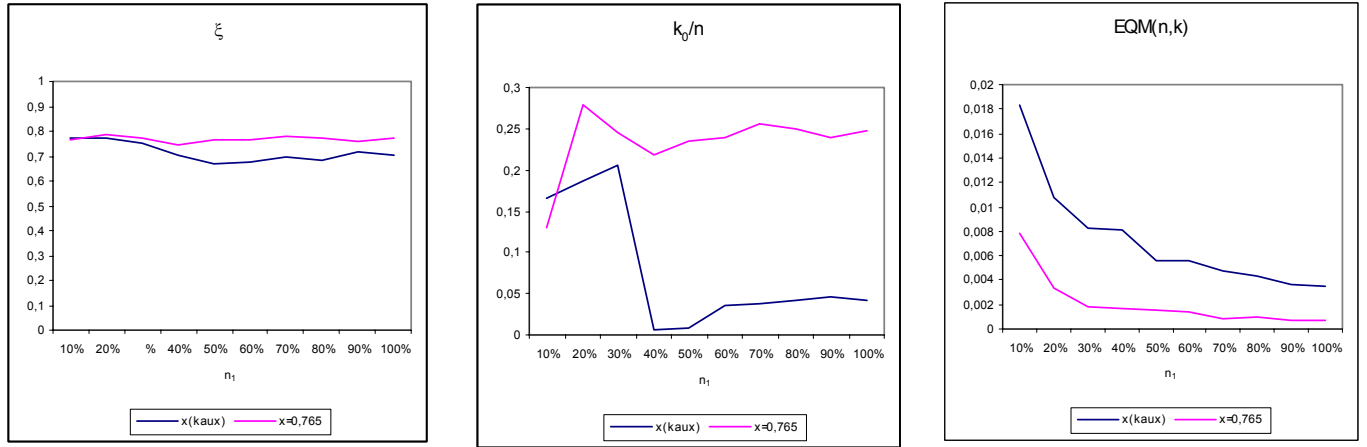


**Figura 3.2.2** – Resultados para diferentes valores de  $n_1$  dado  $\xi_{n,aux} = 0,765$ .

Um dos problemas desta abordagem é a determinação do valor de  $n_1$  já que, como se pode averiguar nos gráficos, diferentes valores de  $n_1$  conduzem a resultados diferentes.

Para ultrapassar este problema, optou-se por escolher um valor de  $n_1$  que minimize o EQM. No gráfico da direita, esse valor corresponde a 90% de  $n$ , de que resulta  $\hat{k}_0(n; \xi_{n,aux}, n_1) = 786$  e  $\hat{\xi} = 0,76177$ .

Outro problema desta mesma abordagem é a forte dependência do valor de  $\xi_{n,aux}$ , como por exemplo, se  $\xi_{n,aux}$  for estimado através de um valor inicial de  $k_{n,aux} = 2\sqrt{n}$ , como apresenta Gomes *et al.* (2002). Os resultados que se obtêm são totalmente diferentes como se pode confirmar nos gráficos da Figura 3.2.3.



**Figura 3.2.3** – Resultados para diferentes valores de  $n_1$  dado diferentes valores de  $\xi_{n,aux}$

Uma vez mais, o EQM é mínimo em  $n_1 = 90\%n$ , de onde resulta  $\hat{k}_0(n; \xi_{n,aux}, n_1) = 155$  e  $\hat{\xi} = 0,71659$ .

No sentido de contornar o problema levantado Danielsson *et al.* (2001) e Gomes *et al.* (2002) sugerem a utilização de estatísticas auxiliares, cujo comportamento assintótico se assemelha ao de  $\xi_n(k) - \xi$ .

Seja

$$M_n(k) = \frac{1}{k} \left( \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^2,$$

a estatística auxiliar proposta por Danielson *et al.* (2001) é

$$T'_n = M_n(k) - 2[\xi_n(k)]^2$$

Assim, em vez de minimizar

$$\text{EQM}(n_1, k_1) = E \left[ \left( \xi_{n_1}^*(k_1) - \xi_{n,aux} \right)^2 \middle| \underline{X}_n \right],$$

minimiza-se

$$\text{EQM}'(n_1, k_1) = E \left[ \left( M_{n_1}(k) - 2[\xi_{n_1}(k)]^2 \right)^2 \middle| \underline{X}_n \right].$$

Seja  $k'_0(n)$  que minimiza  $\text{EQM}'(n, k)$

$$\frac{k_0(n)}{k'_0(n)} = \left( 1 - \frac{1}{\rho} \right)^{-2/(1-2\rho)} (1 + o(1)) \quad (3.4)$$

e

$$\frac{k_0^{'*}(n_1)}{k'_0(n)} = \left( \frac{n_1}{n} \right)^{2\rho/(2\rho-1)} (1 + o(1)).$$

Para estimar  $k'_0(n)$  recorre-se a uma segunda amostra de dimensão  $n_2 = n_1^2 / n$ , devido à seguinte relação:

$$\frac{[k_0^{'*}(n_1)]^2}{k_0^{'*}(n_2)} = k'_0(n)(1 - o(1))$$

à medida que  $n \rightarrow \infty$ .

Em (3.4) constata-se que é necessário estimar  $\rho$  para se obter a estimativa de  $k_0(n)$ .

Os autores do artigo sugeriram o seguinte estimador

$$\rho^{'*} = \frac{\ln \bar{k}_0^{'*}(n_1)}{2 \ln(\bar{k}_0^{'*}(n_1) / n_1)}.$$

Desta forma a estimativa de  $k_0(n; n_1)$  é dada pela fórmula a seguir representada:

$$\hat{k}_0(n; n_1) = \left( 1 - \frac{1}{\rho^{'*}} \right)^{-2/(1-2\rho^{'*})} \left[ \frac{[\bar{k}_0^{'*}(n_1)]^2}{\bar{k}_0^{'*}(n_1^2 / n)} \right].$$

Os resultados que se apresentam resultam do seguinte algoritmo para diferentes valores de  $n_1$ , expressos em percentagem de  $n$  :

1. Retiraram-se 1000 ( $B$ ) amostras *bootstrap*  $(x_1^*, \dots, x_{n_2}^*)$  e  $(x_1^*, \dots, x_{n_2}^*, x_{n_2+1}^*, \dots, x_{n_1}^*)$ ,
2. Para cada valor da amostra foi calculado para cada  $k$  o valor da estatística

$$T'_{n_i}(k) = M_{n_i}(k) - 2[\xi_{n_i}(k)]^2, \quad i = 1, 2$$

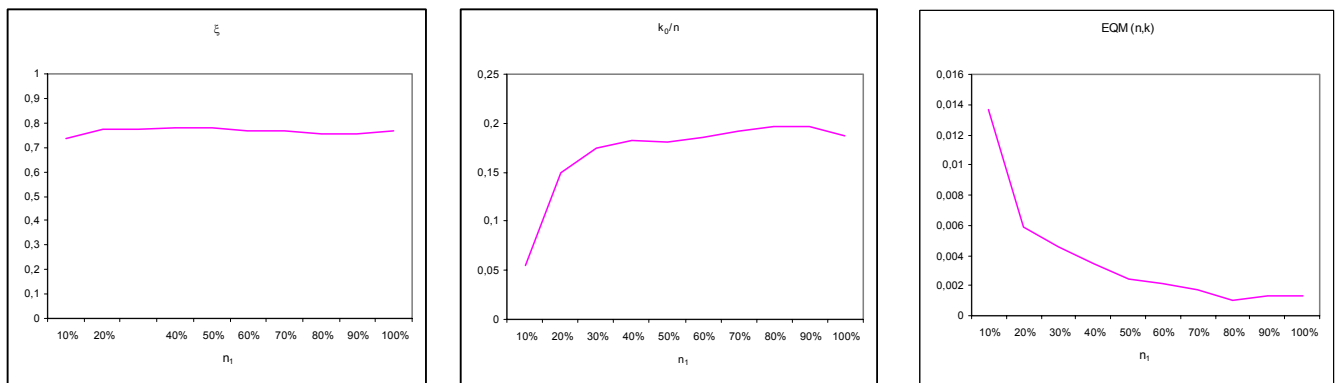
3. Obteve-se  $\bar{k}_0^{'*}(n_i) = \arg \min_k EQM'(n_i, k)$ ,  $i = 1, 2$ , onde

$$EQM'(n_i, k) = \frac{1}{B} \sum_{l=1}^B t'_{n_i, l}(k)^2$$

$$5. \text{ Calculou-se } \hat{k}_0(n; n_1) = \left(1 - \frac{1}{\rho^{'*}}\right)^{-2/(1-2\rho^{'*})} \left[ \frac{[\bar{k}_0^{'*}(n_1)]^2}{\bar{k}_0^{'*}(n_1^2/n)} \right]$$

6. Por fim, obteve-se ao valor de  $\xi_{n, n_1}^{'*} = \xi_n(\hat{k}_0(n; n_1))$ .

Os resultados encontram-se resumidos nos gráficos da Figura 3.2.4.

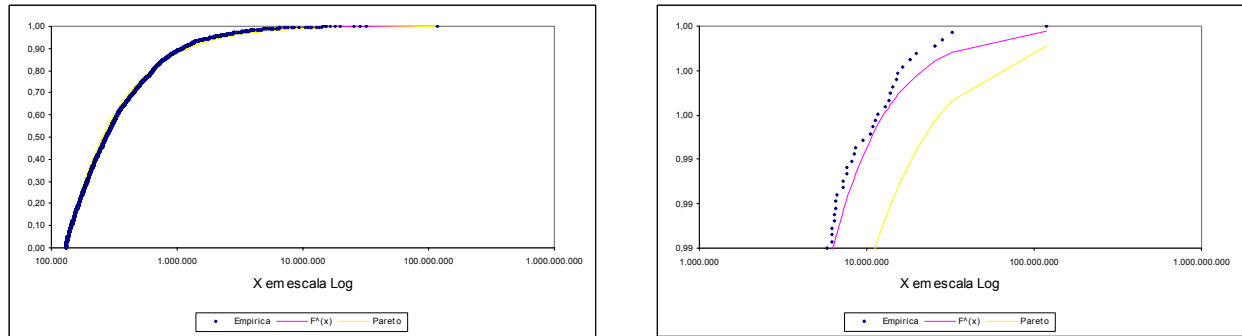


**Figura 3.2.4** – Resultados para diferentes valores de  $n_1$ , segundo a abordagem de Danielsson.



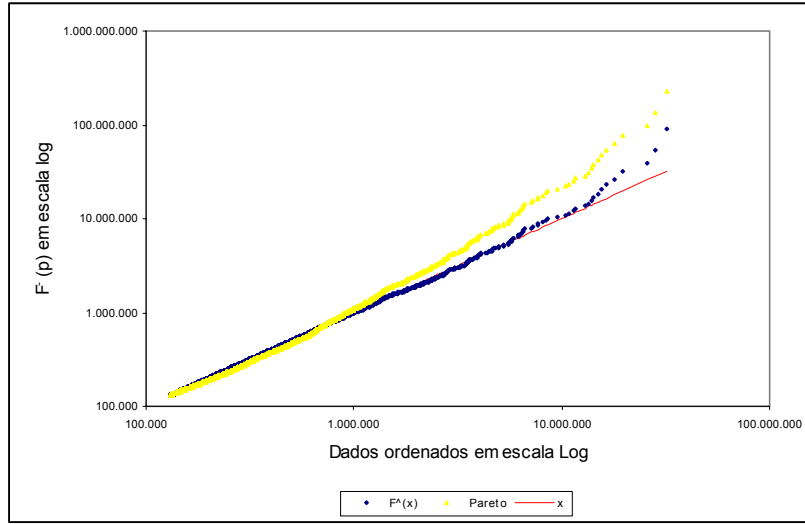
O valor de  $n_1$  que minimiza o EQM, é  $n_1 = 80\%n$ . No gráfico da direita está representado o  $EQM'(n, k)$  para diferentes valores de  $n_1$ , de onde resultam  $\hat{k}_0(n; n_1) = 633$ ,  $\xi_{n, n_1}^* = 0,76467$  e  $x_{k^*, n} = 655.860$ .

Para as indemnizações,  $X$ , inferiores ao nível  $x_{k^*, n}$ , é de considerar a sua modelação pela função de distribuição empírica pelas razões invocadas na Secção 3.1. Na Figura 3.2.5 está representada  $F(x)$  obtida por este método, bem como a distribuição de Pareto do Capítulo 3.



**Figura 3.2.5** – No gráfico da esquerda pode-se observar a distribuição empírica e a distribuição ajustada, segunda a abordagem de Danielsson. O gráfico da direita tem a escala ampliada na cauda.

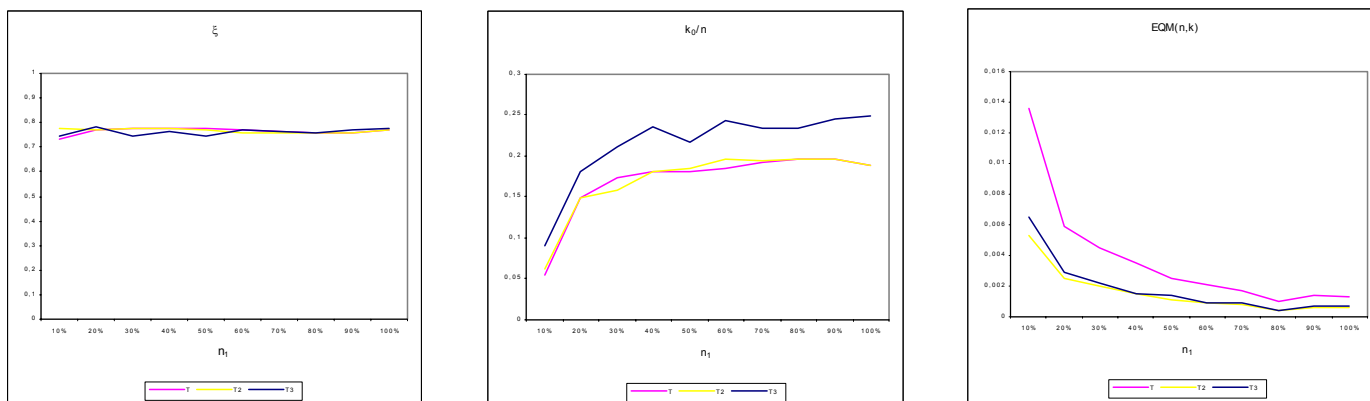
De novo o ajustamento global parece ser aceitável. Ampliando a escala na parte da cauda, verifica-se que existe uma melhoria no ajustamento em comparação com a distribuição de Pareto, apresentada no Capítulo 2. A mesma conclusão pode ser retirada do gráfico Q-Q que se segue.



**Figura 3.2.6** – Comparação dos quantis com os da distribuição ajustada segundo a abordagem de Danielsson.

### 3.3 Discussão de diferentes métodos

Relativamente ao segundo algoritmo apresentado na Secção 3.2.2, importa salientar o facto de os resultados dependerem da estatística auxiliar apresentada. Se a estatística auxiliar  $T'_n = M_n(k) - 2[\xi_n(k)]^2$  for substituída por  $T'_{n,2} = M_n(k) / (2\xi_n(k) - \xi_n(k))$  ou por  $T'_{n,3} = \xi_n(k/2) - \xi_n(k)$ ,  $2 \leq k < n$ , e com as devidas alterações em algumas fórmulas no caso da última estatística auxiliar, como apresentados em Gomes *et al.* (2002), os resultados que se obtêm são diferentes. Os gráficos da Figura 3.3.1 resumizam os resultados apresentados.



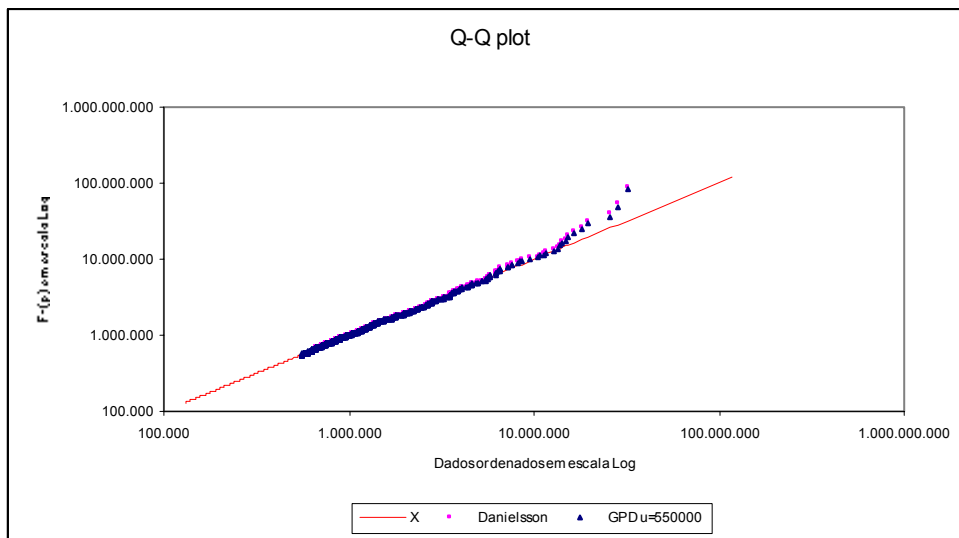
**Figura 3.3.1** – Comparação dos resultados finais de diferentes estatísticas auxiliares.

Existem outros estimadores do parâmetro da condição de primeira ordem alternativos ao utilizado na Secção 3.2. Alguns desses estimadores podem ser encontrados, por exemplo, em Embrechts *et al.* (1997).

Uma grande desvantagem dos métodos adaptativos tem a ver com a morosidade na simulação, facto que condiciona fortemente o número de simulações.

Ao longo do Capítulo foram privilegiados os métodos gráficos no que respeita ao estudo da qualidade do ajustamento, e nenhum teste de hipóteses foi empregue, dado que estes têm alguma dificuldade na detecção das discrepâncias na cauda.

O gráfico seguinte compara os resultados obtidos do modelo da Secção 3.1, com  $u = 550.000$ , com o último modelo apresentado na Secção 3.2.2. Verifica-se que, para os dados em análise, a abordagem semi-paramétrica tem uma cauda ligeiramente mais pesada. Estes dois modelos são utilizados na Secção 4.2 para a modelação das indemnizações agregadas e indemnizações agregadas retidas.



**Figura 3.3.2** –*Comparação dos quantis com os da distribuição ajustada em 3.1 com  $u = 550.000$  e a distribuição ajustada segundo a abordagem de Danielsson.*

## **CAPÍTULO 4**

---

### **Teste de Solvência**

Ao longo dos últimos anos têm surgido no meio actuarial vários modelos sobre a adequabilidade dos capitais das seguradoras por forma a fazer face aos compromissos assumidos pelas empresas daquele sector de actividade económica. A modelação adequada dos quantis elevados da severidade dos sinistros é indispensável para qualquer modelo de solvência.

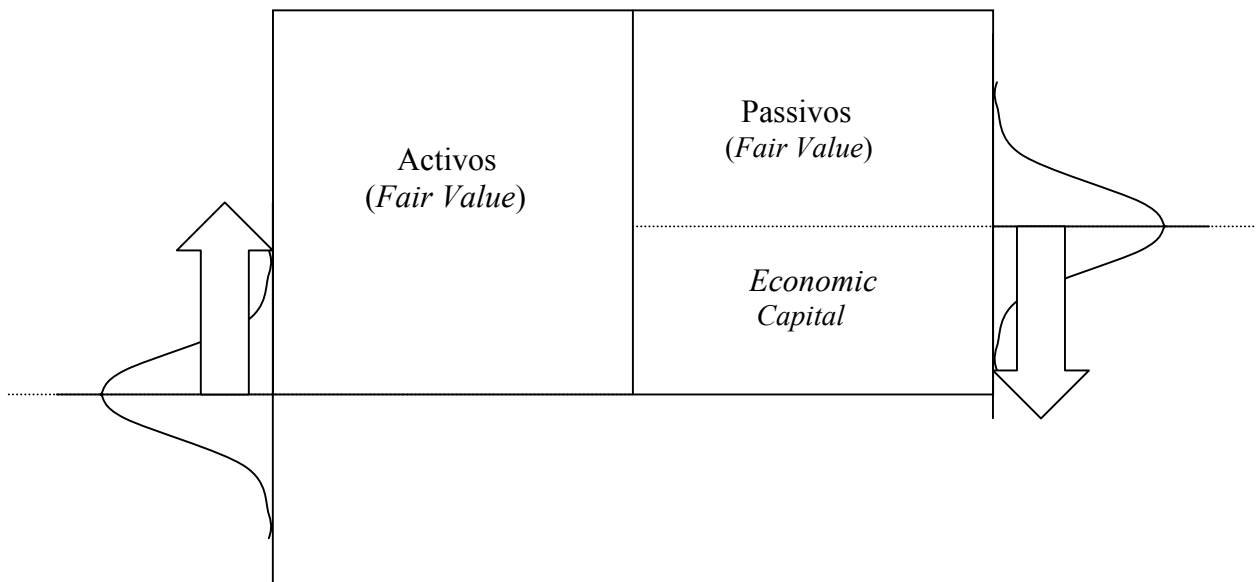
Na Secção 4.1 é exposta, em termos gerais, a ideia base dos Modelos de Solvência e na Secção 4.2 são modeladas as indemnizações agregadas a cargo da Seguradora como módulo de um modelo de solvência, aplicando alguns dos resultados do Capítulo 3. Por fim, na Secção 4.3 é desenvolvida uma breve discussão dos resultados.

#### **4.1 Teste da Solvência**

A ideia central no teste de solvência consiste na avaliação, em termos probabilísticos, da capacidade de uma determinada Companhia de Seguros em responder aos seus compromissos.

Um conceito fortemente ligado ao teste de solvência é *Economic Capital*. Consiste na diferença entre o *fair value* dos activos e o *fair value* das responsabilidades. Por *fair value* entende-se o valor pelo qual um activo pode ser transaccionado ou uma responsabilidade liquidada, entre partes interessadas e bem informadas numa transacção realizável.

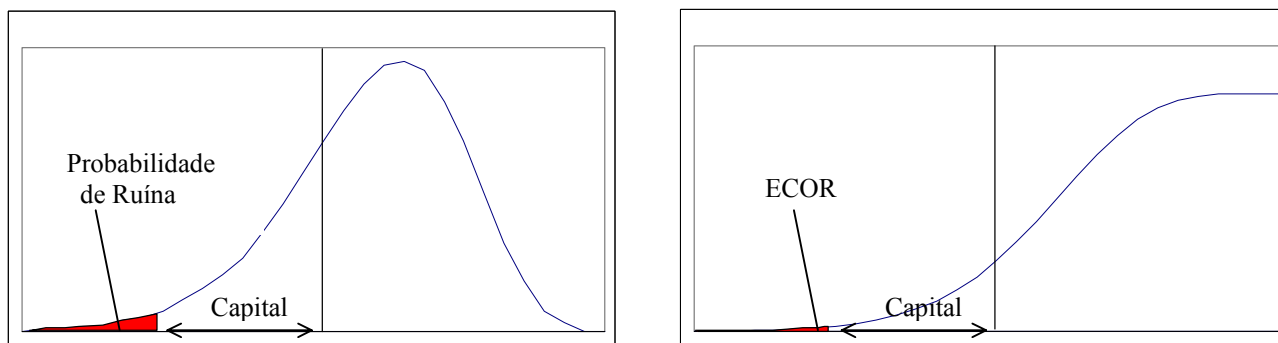
A Figura 4.1.1 esquematiza a relação entre o *Economic Capital* e a solvência.



**Figura 4.1.1** – *Economic Capital em termos esquemáticos.*

O capital mínimo para garantir a solvabilidade é, geralmente, definido em termos de *Probabilidade de Ruína*. Aqui, a probabilidade de ruína é entendida como a probabilidade de insuficiência de capital. Outro conceito ligado à ruína é o *Economic Cost of Ruin* (ECOR) que se traduz no valor esperado da gravidade da ruína.

Os gráficos seguintes sintetizam a ideia subjacente aos dois conceitos anteriormente apresentados.



**Figura 4.1.2** – *Função de Densidade (à esquerda) e Função de Distribuição (à direita) da variação do Economic Capital.*

Existem vários modelos e várias segmentações possíveis dos riscos das Seguradoras. A título de exemplo apresenta-se de seguida a segmentação da agência de *rating* Standard & Poors, onde o risco suportado pelas seguradoras é tipificado do seguinte modo:

- C1 – Investment/ALM Risks (Risco de Crédito, Imunização Inadequada);
- C2 – Reinsurer Credit Risk;
- C3 – Premium Risks (Sinistros futuros, Catastróficos e Não Catastróficos);
- C4 – Reserve Risks (Sinistros Ocorridos);
- C5 – Life Risks (Mortalidade, Calamidade);
- C6 – Residual Risks (ex.: Risco Operacional e Risco de Pequenos Negócios).

Os riscos C3 e C4 são aqueles que têm maior importância em Não Vida.

Embora, por simplicidade, se tenha esquematizado dessa forma os riscos, estes podem não ser independentes entre si, como exemplo, o risco de crédito das resseguradoras não é independente dos riscos catastróficos ou das indemnizações a cargo das resseguradoras.

Outra questão que importa salientar é a dependência entre os rácios de sinistralidade existente entre diferentes linhas de negócios, causada por eventos que conduzem a perdas em diversas linhas de negócio. No caso da dependência se dever a acontecimentos extremos, como por exemplo ataques terroristas ou fenómenos naturais, é preferível que estes sejam modelados à parte. Se a dependência for devido à natureza do risco, é possível que esta seja modelada com funções cópula, de qualquer modo, é importante que se conheça a sua natureza. Em Kreps (2000) é apresentado um algoritmo de simulação de variáveis dependentes. Em Embrechts *et al.* (1999) é feita uma discussão sobre o problema da correlação.

Outro factor que pode favorecer a dependência é o ciclo de subscrição no mercado segurador. Uma forma de contornar esse problema consiste em modelar, em primeiro lugar, as unidades de exposição ao risco e posteriormente, os prémios, de acordo com a estratégia da Companhia ou de acordo com o que se espera do mercado, como em D'Arcy *et al.* (1997) ou em Kaufmann *et al.* (2001), e as indemnizações agregadas associadas à exposição em separado. As indemnizações agregadas são objecto de análise na secção seguinte.

#### **4.2 Modelação das Indemnizações Agregadas**

Apresenta-se nesta Secção uma forma de modelar as indemnizações agregadas retidas, através da aplicação dos resultados obtidos no Capítulo 3. Foram definidos dois modelos, que se distinguem apenas pelo estimador da cauda da distribuição das indemnizações individuais.



Em conformidade com o risco em causa, o programa de resseguro que parece mais adequado é do tipo *excess-of-loss*, devido ao peso das indemnizações extremas nas indemnizações agregadas.

Não foi possível obter outra informação para além do montante individual de cada uma das indemnizações, pelo que, o número de unidade de exposição, a distribuição da frequência dos sinistros e o tipo de tratado de resseguro (o layer e a capacidade) que se apresentam nos modelos que se seguem, são hipóteses que não correspondem à verdadeira carteira em causa.

Seja

$X$  - severidade do sinistro;

$Y$  - severidade líquida de resseguro;

$M$  - número de sinistros de um segurado;

$N$  - número total de sinistros;

$S^{(i)}$  - indemnização agregada do modelo  $i$ ;

$\hat{S}^{(i)}$  - indemnizações agregadas retidas do modelo  $i$ ;

### **Modelo 1**

$$\left. \begin{array}{l} M \sim Poi(0,01) \\ Exposição : 50.000 \end{array} \right\} N \sim Poi(500);$$

$$F(x) = \begin{cases} F_n(x) & x < u \\ 1 - \left( 1 + \xi \frac{x - u - \hat{v}}{\hat{\beta}'} \right)^{-1/\xi} & x \geq u \end{cases}$$

$$\text{onde } \hat{v} = \frac{\hat{\beta}}{\xi} \left( \left( \frac{N_u}{n} \right)^{\xi} - 1 \right) \text{ e } \hat{\beta}' = \hat{\beta} \left( \frac{N_u}{n} \right)^{\xi}.$$

Tipo de resseguro: 985.000.000 xs 15.000.000

### Modelo 2

$$\left. \begin{array}{l} M \sim Poi(0,01) \\ Exposição : 50.000 \end{array} \right\} N \sim Poi(500);$$

$$F(x) = \begin{cases} F_n(x) & x < x_{\hat{k}^*,n} \\ 1 - \frac{k}{n} \left( \frac{x}{X_{k,n}} \right)^{-\hat{\alpha}_{k,n}^{(H)}} & x \geq x_{\hat{k}^*,n} \end{cases}$$

Tipo de resseguro: 985.000.000 xs 15.000.000

Por  $B$  xs  $A$ , entende-se um tratado *excess-of-loss* em que a resseguradora paga até  $B$  por excesso de  $A$ .

Seja  $A_1^{(1)} = \{x : x < u\}$  e  $A_2^{(1)} = \{x : x \geq u\}$ , no caso do modelo 1, e  $A_1^{(2)} = \{x : x < x_{\hat{k}^*,n}\}$  e  $A_2^{(2)} = \{x : x \geq x_{\hat{k}^*,n}\}$ , no caso do modelo 2.

Dada a especificidade da distribuição das indemnizações individuais, em termos operacionais é vantajoso a desagregação das indemnizações agregadas em indemnizações da classe  $A_1^{(i)}$  e em indemnizações da classe  $A_2^{(i)}$ ,  $i = 1, 2$ .

Seja  $u$ , um nível genérico e  $\varphi_s(\omega)$ , a função característica de  $S$ ,

$$\varphi_s(\omega) = E[e^{i\omega S}] = E[E[e^{i\omega S} | N]] = E[\varphi_X(\omega)^N] = E[e^{N \ln \varphi_X(\omega)}] = M_N[\ln \varphi_X(\omega)]$$

em que a função característica de  $X$  é

$$\varphi_X(\omega) = E[e^{i\omega X}] = \int_0^\infty e^{i\omega x} dF(x) = \int_0^u e^{i\omega x} dF_{x|x < u}(x) F_n(u) + \int_u^\infty e^{i\omega x} dF_{x|x \geq u}(x) (1 - F_n(u))$$

Seja

$$N_1 \sim Poi(\lambda_1), \lambda_1 = \lambda F_n(u)$$

$$N_2 \sim Poi(\lambda_2), \lambda_2 = \lambda(1 - F_n(u))$$

$$\varphi_{X_1}(\omega) = \int_0^u e^{i\omega x} dF_{x|X < u}(x)$$

$$\varphi_{X_2}(\omega) = \int_u^\infty e^{i\omega x} dF_{x|X \geq u}(x)$$

onde  $X_1$  é a v.a.  $X$  dado que é inferior a  $u$  e  $X_2$  dado que é superior ou igual a  $u$ .

$\varphi_s(\omega)$  pode se rescrito da seguinte forma

$$\varphi_s(\omega) = \exp\{\lambda[\varphi_{X_1}(\omega)F_n(u) + \varphi_{X_2}(\omega)(1 - F_n(u)) - 1]\} = \prod_{i=1}^2 \exp\{\lambda_i[\varphi_{X_i}(\omega) - 1]\} = \varphi_{S_1+S_2}(\omega)$$

onde,  $S_1$  são indenizações agregadas de indenizações pertencentes à classe  $A_1$  e  $S_2$  as indenizações agregadas de indenizações pertencentes à classe  $A_2$ .

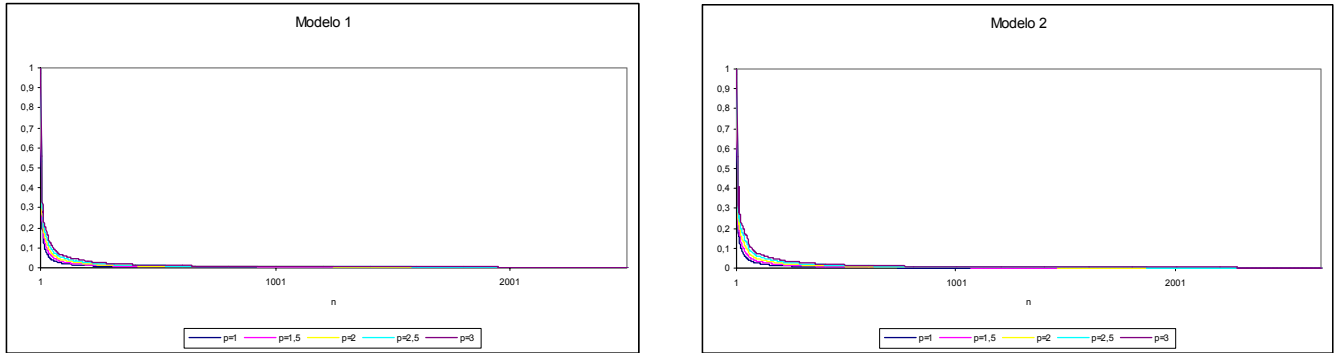
Ou seja, por simplicidade é possível modelar as indenizações  $S_1$  e  $S_2$  em separado com ligeira alteração dos parâmetros.

*Modelação de  $S_1^{(1)}$  e  $S_1^{(2)}$*

Ao longo dos capítulos anteriores, a análise tem sido centrada nas indenizações de elevado montante. Para a modelação de  $S_1^{(1)}$  e  $S_1^{(2)}$  é necessário que se conheça melhor a natureza das indenizações individuais de pequeno montante.

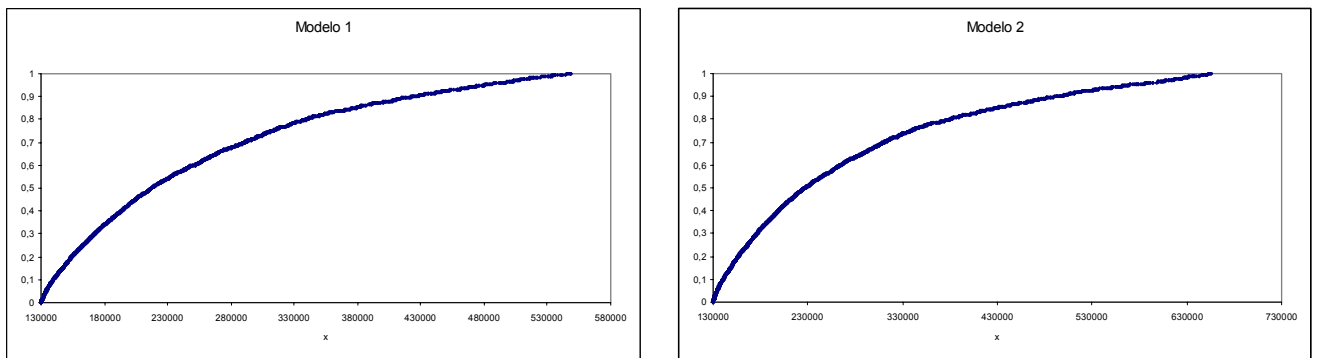
Uma vez retiradas as indenizações de classe  $A_2$  dos dados, prevê-se, por um lado, que as indenizações se tornem num grupo mais homogêneo e que os valores observados sejam contíguos entre si. Espera-se, por outro lado, que o impacto de uma indenização individual seja negligenciável na soma.

Na Figura 4.2.1 encontra-se representado  $R_{n_i}(p)$  com diferentes valores de  $p$  para ambos os modelos. Pode-se verificar que, em qualquer um dos modelos,  $R_{n_i}(p)$  tende rapidamente para zero, indicando que o efeito do máximo na soma é quase nulo.



**Figura 4.2.1** –  $R_n(p)$  das indenizações da classe  $A_1$  do Modelo 1 e Modelo 2.

Os gráficos seguintes representam a função de distribuição empírica das indenizações da classe  $A_1$  de cada um dos modelos. Em ambos os modelos constata-se a existência de uma certa contiguidade entre os diferentes valores observados.



**Figura 4.2.2** – Função de distribuição empírica das indenizações da classe  $A_1$  do Modelo 1 e Modelo 2.

A simulação do valor de cada uma das indenizações individuais foi feita por reamostragem com reposição. Nesta classe de indenizações, não é modelado o

resseguro dado que em ambos os modelos, nenhuma indemnização da classe  $A_1$  excede o *layer* definido para o resseguro.

*Modelação da distribuição de  $S_2^{(1)}$ ,  $S_2^{(2)}$ ,  $\hat{S}_2^{(1)}$  e  $\hat{S}_2^{(2)}$*

Novamente as distribuições de  $S_2^{(1)}$ ,  $S_2^{(2)}$ ,  $\hat{S}_2^{(1)}$  e  $\hat{S}_2^{(2)}$  foram obtidas por simulação.

Os valores da severidade da Classe  $A_2$  no Modelo 1, foram simulados através da seguinte expressão:

$$X_2 = u + \frac{\hat{\beta}}{\hat{\xi}} \left( (1 - P)^{-\hat{\xi}} - 1 \right), \quad P \sim U(0,1)$$

ou seja, foram simulados os excessos em relação a  $u$ , com a distribuição condicionada obtida no capítulo anterior.

No Modelo 2, a expressão utilizada foi a seguinte:

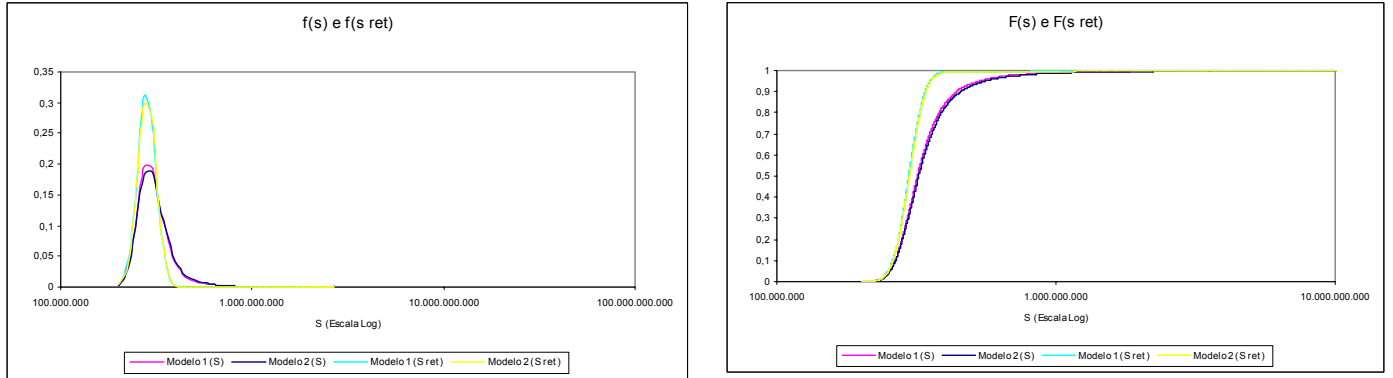
$$X_2 = \left( \frac{n}{\hat{k}^*} \left( 1 - F(x_{\hat{k}^*,n}) - (1 - F(x_{\hat{k}^*,n}))P \right) \right)^{-\frac{\hat{\xi}}{\xi}} x_{\hat{k}^*,n}, \quad P \sim U(0,1).$$

Ao contrário do que aconteceu nas indemnizações da classe  $A_1$ , na presente classe é necessário modelar as indemnizações retidas. Representam-se por  $\hat{S}_2^{(1)}$  e  $\hat{S}_2^{(2)}$  as indemnizações agregadas retidas das indemnizações das classes  $A_2^{(1)}$  e  $A_2^{(2)}$  respectivamente.

*Distribuição de  $S^{(1)}$ ,  $S^{(2)}$ ,  $\hat{S}^{(1)}$  e  $\hat{S}^{(2)}$*

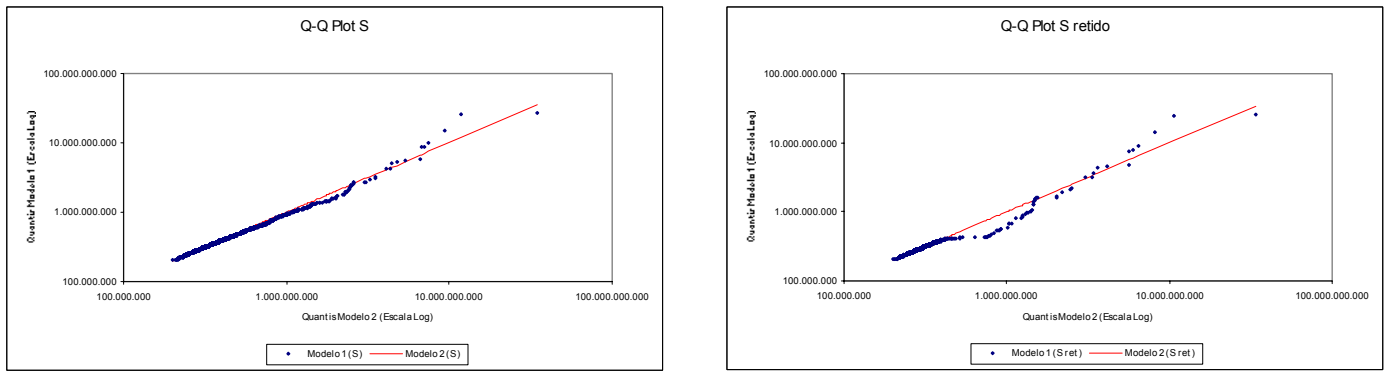
No Modelo 1, adicionando os valores simulados de  $S_1^{(1)}$  a  $S_2^{(1)}$  e a  $\hat{S}_2^{(1)}$ , obtêm-se os resultados das simulações de  $S^{(1)}$  e  $\hat{S}^{(1)}$ , respectivamente. No Modelo 2, os resultados podem ser obtidos de modo semelhante.

Na Figura 4.2.4 encontram-se representadas as funções de densidade e de distribuição de  $S^{(1)}$ ,  $S^{(2)}$ ,  $\hat{S}^{(1)}$  e de  $\hat{S}^{(2)}$ .



**Figura 4.2.3** – Funções de densidade (à esquerda) e de distribuição (à direita) de  $S^{(1)}$ ,  $S^{(2)}$ ,  $\hat{S}^{(1)}$  e  $\hat{S}^{(2)}$ .

A Figura 4.2.3 indicia que o Modelo 1 e o Modelo 2 produzem resultados idênticos, especialmente após o resseguro. Porém, através da análise dos gráficos Q-Q da Figura 4.2.4, constata-se que os resultados divergem nos quantis elevados. Esta divergência deve-se, sobretudo, às indemnizações que excedem a capacidade do tratado, i.e., indemnizações superiores a €1.000.000.000. Isto resulta do facto de no Modelo 2 a distribuição da severidade ter uma cauda ligeiramente mais pesada do que a distribuição da severidade no Modelo 1, como verificado na Secção 4.3.



**Figura 4.2.4** – Comparação dos quantis estimados de  $S^{(1)}$ ,  $S^{(2)}$ ,  $\hat{S}^{(1)}$  e  $\hat{S}^{(2)}$ .

No entanto, o máximo observado foi €117.953.980 que, por sua vez, é 3,66 vezes maior do que a estatística de segunda ordem. De imediato levantam-se as seguintes questões:

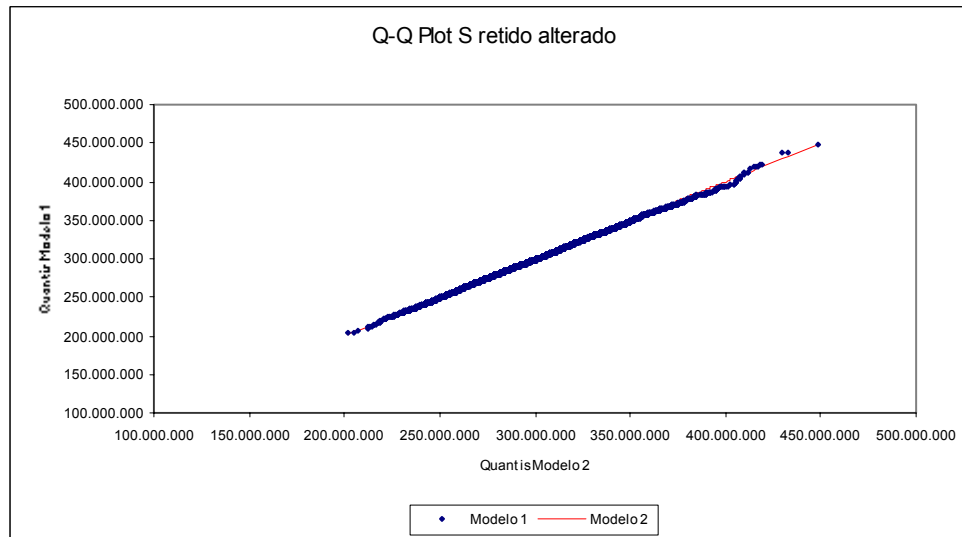
*Será aceitável projectar quantis dez ou vinte vezes superiores ao máximo observado?*

*Será esse valor projectado coerente com os dados da carteira da Seguradora?*

Apesar de as técnicas apresentadas no Capítulo 3 possibilitarem a projecção de quantis superiores ao máximo observado, a sua precisão é menor à medida que o quantil projectado se afasta do máximo observado. O conhecimento do *Maximum Possible Loss* (MPL) de uma determinada carteira de seguros, permite limitar a variabilidade dos quantis projectados, como por exemplo, censurando os valores à sua direita, i.e.,  $\min\{X, MPL\}$ .

Dado que não se considera correcta a estimação de quantis demasiados elevados através dos estimadores de cauda apresentados, nomeadamente a estimação de indemnizações que excedem a capacidade do tratado, poderá não fazer sentido, neste tipo de análise, a modelação de tratados de resseguro com capacidade limitada. Em vez disso, talvez seja preferível a modelação de tratados com capacidade ilimitada. O impacto de sinistros que excedem a capacidade podem ser modelados, à parte, como eventos de um em  $x$  anos.

Os quantis das indemnizações retidas para um tratado com capacidade ilimitada estão apresentados na Figura 4.2.5, e verifica-se que o Modelo 1 e o Modelo 2 conduzem a resultados semelhantes.



**Figura 4.2.5** – *Comparação dos quantis das indemnizações agregadas retidas estimados nos dois modelos considerando ilimitada a capacidade do tratado.*

### 4.3 Discussão

A modelação da dependência entre diferentes linhas de negócio tem sido objecto de debate entre os actuários. Sendo a modelação por funções cópula a abordagem mais comum. No entanto, devido às possíveis alterações das condições contratuais, devem-se utilizar com algum cuidado os dados históricos para se proceder ao estudo da dependência. Em último caso é possível recorrer a procedimentos de estudo da sensibilidade. Informação sobre funções cópula pode ser encontrada em Nelsen (1999).

Outro tópico de enorme debate é o ciclo de subscrição, alguns actuários preferem modelar directamente os rácios de sinistralidade, contornando o problema do ciclo do



mercado com séries temporais, como alternativa ao uso das cadeias de Markov em tempo discreto que é sugerido no texto.

O lado esquerdo da expressão  $\lim_{n \uparrow \infty} n\bar{F}(c_n x + d_n) = -\ln \Phi_\xi(x)$ , já apresentado no segundo Capítulo, pode ser interpretada como o número de observações normalizados da sequência  $(X_1 - d_n)/c_n, \dots, (X_n - d_n)/c_n$  que excedem o nível  $x$ . Esse número de observações segue uma Binomial com parâmetros  $(n, \bar{F}(c_n x + d_n))$ . À medida que  $n \rightarrow \infty$  e  $x \rightarrow \infty$ , a distribuição pode ser aproximada com uma Poisson.

A desagregação das indenizações agregadas pode estender-se a outras distribuições de frequência da classe  $(a, b, 0)$ . Seja  $p_1 = \text{Prob}(X \in A_1)$  e  $p_2 = \text{Prob}(X \in A_2)$ . Para o caso da Binomial Negativa, com função de probabilidade

$$p_k = \binom{\alpha + k - 1}{k} \left( \frac{\beta}{\beta + 1} \right)^\alpha \left( \frac{1}{\beta + 1} \right)^k$$

a desagregação pode ser obtida por alteração do parâmetro  $\beta$ , sendo o parâmetro do número das indenização da classe  $A_1$   $\beta / p_1$ , e da classe  $A_2$   $\beta / p_2$ . Como  $\lim_{p_2 \uparrow 0} [(\beta / p_2) / (\beta / p_2 + 1)] = 1$ , para  $\alpha \rightarrow \infty$ , a distribuição da frequência tende para uma distribuição Poisson, o que é consistente com o resultado apresentado em cima.

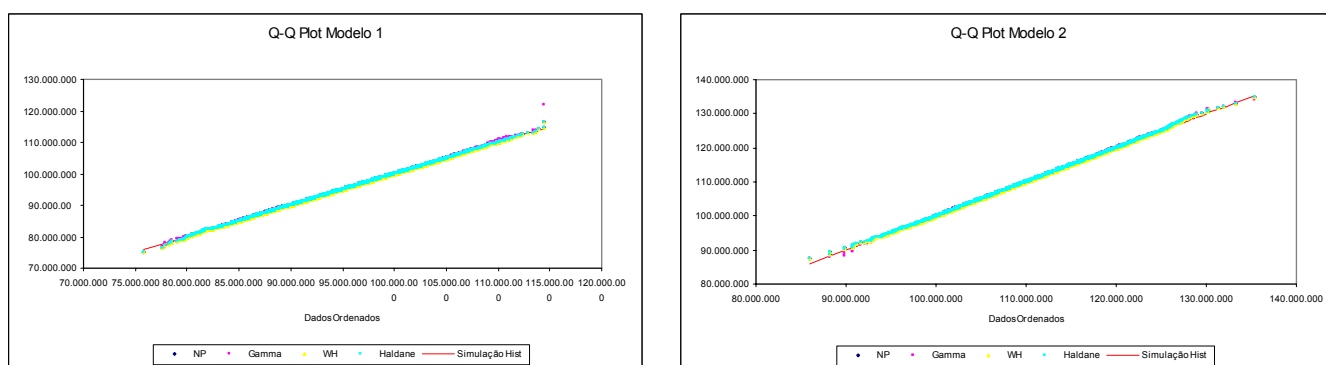
Para uma Binomial com parâmetros  $(n, p)$ , a desagregação é obtida modificando o parâmetro  $p$ , sendo os parâmetros  $p_1 p$  e  $p_2 p$  para a distribuição da frequência dos sinistros das classes  $A_1$  e  $A_2$ , respectivamente. Novamente,  $\lim_{p_2 \uparrow 0} p_2 p = 0$ , para  $n \rightarrow \infty$ , a distribuição tende para uma Poisson.

Relativamente à modelação de  $S_1^{(1)}$  e  $S_1^{(2)}$ , a simulação histórica pode-se tornar num processo consumidor de tempo, sobretudo quando se tratam de carteiras onde o número

esperado de sinistros atinge valores consideráveis. Este processo pode ser simplificado simulando  $S_1^{(1)}$  e  $S_1^{(2)}$  em termos agregados.

À medida que número esperado de sinistros aumenta,  $S_1^{(1)}$  e  $S_1^{(2)}$  tendem para a normalidade. Para valores intermédios do número de sinistros, espera-se que  $S_1^{(1)}$  e  $S_1^{(2)}$  sejam ligeiramente assimétricos. Se isto for verdade, os métodos de aproximação mais usuais, como a Normal-Power; a aproximação à Gamma deslocada; a fórmula de Wilson-Hilferty ou a transformação A da abordagem de Haldane conseguem fornecer uma aproximação aceitável para a distribuição de  $S_1^{(1)}$  e  $S_1^{(2)}$ . Estes métodos de aproximação tornam-se particularmente úteis, quando o número esperado de sinistros atinge valores elevados que inviabilizam o recurso à simulação histórica. Os coeficientes de assimetria de  $S_1^{(1)}$  e  $S_1^{(2)}$  são 0,06490 e 0,06699, respectivamente.

Os gráficos da Figura 4.3.1 comparam os resultados de dez mil simulações por reamostragem de  $S_1^{(1)}$  e de  $S_1^{(2)}$  com os quantis dos métodos de aproximação enunciados. É possível observar que não existem diferenças significativas entre os métodos de aproximação e os resultados da simulação. Reconhece-se, contudo, que os valores obtidos através de reamostragem, são em si uma aproximação aos quantis de  $S_1^{(1)}$  e  $S_1^{(2)}$ , não significando que esta é a exacta.



**Figura 4.3.1** – Comparação dos quantis dos valores simulados de  $S_1^{(1)}$  e de  $S_1^{(2)}$  e os quantis dos métodos aproximados.

No entanto, é difícil generalizar a ideia de que os métodos de aproximação garantem bons resultados em todas as situações. A qualidade das aproximações nem sempre é verificável e dependem sobretudo dos dados em causa.

As fórmulas e os parâmetros dos modelos de aproximação referidos encontram-se no Anexo 1. Informação mais detalhada sobre os métodos de aproximação das indemnizações agregadas pode ser encontrado em Daykin *et al.* (1994), Centeno (2001), Beard *et al.* (1984). A aproximação Normal-Power pode, em alguns casos, subestimar a cauda, o que pode por em causa o resultado de todo um trabalho cuja preocupação principal foi a estimação dos quantis superiores das indemnizações agregadas. Ainda assim, em Pentikäinen (1977) são apresentados alguns argumentos sobre a superioridade da aproximação NP em relação à Gamma deslocada. Uma discussão sobre a qualidade do ajustamento da fórmula de Wilson-Hilferty e da transformação A da abordagem de Haldane, pode ser encontrada em Pentikäinen (1987).

No sentido de confirmar as hipóteses assumidas nos dois modelos, é possível confirmar o valor esperado das indemnizações a cargo do Ressegurador e o número de sinistros

declarados, com a experiência passada da Companhia. O valor do prémio puro por *layer* é calculado com a seguinte fórmula

$$\Pi(r, R) = \int_r^R (x - r) dF(x) + \bar{F}(R)(R - r) = \int_r^R \bar{F}(x) dx$$

onde  $r$  é a prioridade e  $R - r$ , a capacidade do tratado. De onde resultam, 99.658 no Modelo 1 e 113.315 no Modelo 2.

O número esperado de sinistros a declarar ao Ressegurador, dado por  $\lambda \bar{F}(r)$ , são no Modelo 1 e no Modelo 2, 1,50421 e 1,60245, respectivamente.

O valor esperado das indemnizações agregadas cedidas são 49.829.000 no Modelo 1 e 56.657.500 no Modelo 2. O valor é superior no Modelo 2, devido ao facto de a cauda das indemnizações individuais ser mais pesada. Ainda que os quantis estimados pelos dois modelos sejam semelhantes como verificado na Figura 3.3.1. A modelação das indemnizações agregadas cedidas é importante também para a tomada de decisão.

## CAPÍTULO 5

---

### Conclusão

Ao longo deste trabalho tentou-se demonstrar a utilidade da Teoria de Valores Extremos, na actividade seguradora, no que respeita à análise da cauda das distribuições, onde a abordagem a que os actuários estão habituados pode sub ou sobre estimar a cauda e onde a simulação histórica se revela insuficiente na análise de eventos extremos.

Embora a natureza dos dados analisados seja de cauda pesada, as abordagens apresentadas no Capítulo 3 conseguem oferecer um bom ajustamento em dados provenientes de diversos ramos da actividade seguradora. Há, no entanto, que ter em conta o número de observações consideradas para estimar a cauda, já que, se este for reduzido, poderá conduzir a um maior desvio padrão associado ao parâmetro de cauda.

As abordagens apresentadas encontram-se ainda em fase de desenvolvimento. No caso da distribuição de Pareto Generalizada, em diversas situações o nível óptimo de  $u$  não é claro, a emergência de um algoritmo adaptativo pode contribuir para a sua determinação. Já no que se refere à abordagem semi-paramétrica os métodos adaptativos têm sido um dos temas de intensa investigação, ao longo dos últimos anos.

Na análise aqui considerada, assumiu-se que as indemnizações eram independentes e identicamente distribuídas. Na realidade essa hipótese pode não se verificar. Em algumas situações as indemnizações podem ter uma tendência crescente como se apresenta em McNeil (1998).

Outra questão que importa salientar relaciona-se com os dados. Ao tentar obter o maior número de observações de eventos extremos, o analista necessita muitas vezes de recorrer a dados relativamente antigos. Ao fazê-lo poderá estar a trabalhar com dados que pouco têm a ver com o perfil actual da carteira e/ou do risco em causa. O contrário, ao utilizar dados demasiado recentes, enfrenta o problema do tempo de encerramento de sinistros extremos, que pode ser longo, deparando-se com um elevado grau de incerteza sobre o custo final, resultante da falta de experiência passada. Como resultado, alguns valores arredondados, como 10.000.000 ou 150.000.000, podem pesar no conjunto de indemnizações de elevado montante, perturbando a análise.

## ANEXO I

---

### A.1. Aproximação *Normal-Power*

$$F(s) \approx \Phi \left( -\frac{3}{\gamma_s} + \sqrt{\frac{9}{\gamma_s^2} + 1} + \frac{6}{\gamma_s} \frac{s - \mu_s}{\sigma_s} \right),$$

onde

$$\mu_s = E[S] = E[N]E[X],$$

$$\sigma_s = \sqrt{E[N]V[X] + V[N]E^2[X]},$$

$$\gamma_s = \frac{\mu_3[N]E^3[X] + 3V[N]E[X]V[X] + E[N]\mu_3[X]}{\sigma_s^3}.$$

Na aproximação *Normal-Power* os parâmetros foram os seguintes:

	Modelo 1	Modelo 2
$\mu_{S_1}$	94975337,97	110135199,85
$\sigma_{S_1}$	5307777,20	6113831,51
$\gamma_{S_1}$	0,06490	0,06699

## A.2. Aproximação à *Gamma Deslocada*

Seja

$$S = k + Y$$

em que  $f(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta y} y^{\alpha-1}$ ,  $y > 0$ . Os parâmetros  $\alpha$ ,  $\beta$  e  $k$  são obtidos resolvendo o

seguinte sistema de equações

$$\mu_{S_1} = k + \frac{\alpha}{\beta}, \sigma_{S_1}^2 = \frac{\alpha}{\beta^2}, \gamma_{S_1} = \frac{2}{\sqrt{\alpha}}.$$

Na aproximação à *Gamma Deslocada* os parâmetros foram os seguintes:

	Modelo 1	Modelo 2
$\alpha$	949,74	891,22
$\beta$	0,000058062	0,000048830
$k$	-68598718,79	-72382384,92

## A.3. Fórmula *Wilson-Hilferty*

$$F(s) \approx \Phi \left( c_1 + c_2 \left( \frac{s - \mu_{S_1}}{\sigma_{S_1}} + c_3 \right)^{1/3} \right),$$

onde

$$c_1 = \frac{1}{3g} - 3g, c_2 = 3g^{2/3}, c_3 = g \text{ e } g = \frac{2}{\gamma_{S_1}}.$$

A fórmula *Wilson-Hilferty* é um caso particular da abordagem de Haldane, com

$h = 1/3$ , sendo também uma aproximação à distribuição *Gamma*.



Na fórmula *Wilson-Hilferty* os parâmetros foram os seguintes:

	Modelo 1	Modelo 2
$c_1$	-92,44261	-89,54851
$c_2$	29,48871	28,87013
$c_3$	30,81781	29,85322

#### A.4. Transformação A da Abordagem de Haldane

Seja

$$Y = (S / \mu_S)^h \text{ e } y_{norm} = \frac{(S / \mu_S)^h - \mu_Y}{\sigma_Y},$$

e o seguinte parâmetro auxiliar

$$r = \sigma_S / \mu_S$$

os parâmetros da abordagem de Haldane, podem ser escritos da seguinte forma

$$h = 1 - \frac{1}{3} \gamma_S / r,$$

$$\mu_Y = 1 - \frac{1}{2} h(1-h) \left[ 1 - \frac{1}{4} (2-h)(1-3h)r^2 \right] r^2,$$

$$\sigma_Y = hr \sqrt{1 - \frac{1}{2} (1-h)(1-3h)r^2},$$

$$y_{norm} = \left[ \left( 1 + \frac{S - \mu_S}{\sigma_S} r \right)^h - \mu_Y \right] / \sigma_Y,$$

$$F(s) \approx \Phi(y_{norm}).$$

Na transformação A da abordagem de Haldane os parâmetros foram os que abaixo se indicam:

	Modelo 1	Modelo 2
$r$	0,05589	0,05551
$h$	0,61292	0,59772
$\mu_Y$	0,99963	0,99963
$\sigma_Y$	0,03426	0,03319

## BIBLIOGRAFIA

---

BEARD, R.E., PENTIKÄINEN, T. AND PESONEN, M. (1984) *Risk Theory – The Stochastic Basis of Insurance*. 3<sup>rd</sup> Ed. Chapman & Hall.

BEIRLANT, J., MATTHYS, G. AND DIERCKX, G. (2001) Heavy-Tailed Distributions and Rating. *ASTIN Bulletin* 31, 37-58.

CANTO E CASTRO, L.C. (1992) *Sobre a Teoria Assintótica de Extremos*. Tese de Doutorado. C.E.A.U.L.

CASELLA, G. AND BERGER, R.L. (1990) *Statistical Inference*. Wadsworth & Brooks, Pacific Grove, California.

CENTENO, M. L. (2003), *Teoria do Risco na Actividade Seguradora*. Celta Editora. Colecção Económicas. Oeiras.

CHAMBERS, J.M., CEVELAND, W.S., KLEINER, B. AND TUKEY, P.A. (1983) *Graphical Methods for Data Analysis*. Wadsworth & Brooks, Pacific Grove, California.

D'ARCY, S.P., GORVETT, R.W., HERBERS, J.A., HETTINGER, T.E., LEHMANN, S.G. AND MILLER, M.J. (1997) Building a Public Access PC-Based DFA Model. *Casualty Actuarial Forum*, 1-40.

DANIELSSON, J., HAAN, L. DE, PENG, L. AND VRIES, C.G. DE (2000) Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation. *Econometric Institute Report* EI 2000-19/A.

DANIELSON, J. AND VRIES, C.G. DE (1997) Beyond the Sample: Extreme Quantile and Probability Estimation. Disponível em <http://www.hag.hi.is/~jond/research>.

DAYKIN, C.D., PENTIKÄINEN, T. AND PESONEN, M. (1994) *Practical Risk Theory for Actuaries*. Chapman & Hall.

EFRON, B. AND TIBSHIRANI, R.J. (1993) *An Introduction to the Bootstrap*. Chapman & Hall/CRC

EMBRECHTS, P., KLÜPELBERG, C. AND MIKOSCH, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer.

EMBRECHTS, P., MCNEIL, A. AND STRAUMANN, D. (1999) Correlations: Pitfalls and Alternatives, retirado de <http://www.math.ethz.ch/~embretchs>.

FRAGA ALVES, M.I. (2001) Caudas Gordas ou Pesadas – Como Manusear? *Notas e Comunicações C.E.A.U.L.* 2/2001.

FRAGA ALVES, M.I. (2002) Estimation of First and Second Parameters in Heavy Tails. *Notas e Comunicações C.E.A.U.L.* 7/2002.

GALAMBOS, J. (1987) *The Asymptotic Theory of Extreme Order Statistics*, 2<sup>nd</sup> Ed. Krieger, Malabar, Florida.

GOMES, M.I. (1994) Metodologias Jackknife e Bootstrap em Estatísticas de Extremos, in: Mendes Lopes et al. (ed.), *Actas do II Congresso Anual da Sociedade Portuguesa de Estatística*, 31-46.

GOMES, M.I. (1999) The Jackknife and the Bootstrap Methodologies in the Estimation of Parameters of Rare Events. *Revista de Estatística* 99-I, 7-23.

GOMES, M.I. AND OLIVEIRA, O. (2001) The Bootstrap Methodology in Statistics of Extremes – Choice of the Optimal Sample Fraction. *Extremes* 4, 331-358.

HALL, P. (1982) On Some Simple Estimates of an Exponent of Regular Variation, *J. Roy. Statist. Soc. B*44, 37-42.

HALL, P. (1990) Using the Bootstrap to Estimate Mean Squared Error and Select Smoothing Parameter in Nonparametric Problems. *Journal of Multivariate Analysis* 32, 177-203.

HILL, B.M. (1975) A Simple General Approach to Inference About the Tail of a Distribution, *Ann. Statist.* 3, 1163-1174.

HOGG, R.V. AND KLUGMAN, S.A. (1984) *Loss Distributions*. Wiley, New York.

KAUFMANN, R., GADMER, A. AND KLETT, R. (2001) Introduction to Dynamic Financial Analysis. *ASTIN Bulletin* 31, 213-249.

KREPS, R.E. (2000) A Partially Comonotonic Algorithm for Loss Generation. *Proceeding of XXXIst International ASTIN Colloquium*, 165-176, Porto Cervo, Italy.

MCNEIL, A. (1997) Estimating the Tails of Toss Severity Distributions using Extreme Value Theory. *ASTIN Bulletin* 27, 117-137.

MCNEIL, A. and Saladin, T. (1998) Developing Scenarios for Future Extreme Losses Using a POT Model, retirado de <http://www.math.ethz.ch/~mcneil>.

MURTEIRA, B.J.F. (1998) *Probabilidades e Estatística*, vol. I e II, 2ª Ed. McGraw-Hill.

NELSEN, R.B.(1999) *An Introduction to Copulas*. Springer.

OLIVEIRA, J.T. DE (1997) *Statistical Analysis of Extremes*. Pendor.

REISS, R.D. AND THOMAS, M. (1997) *Statistical Analysis of Extreme Values from Insurance, Finance, Hydrology and other fields*. Birkhäuser Verlag.

RESNICK, S.I. (1997) Discussion of the Danish data on Large Fire Insurance Losses. *ASTIN Bulletin* 27, 139-151.

RESNICK, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes*. Springer-Verlag.

ROSS, M.S. (2002) *Simulation*. 3<sup>rd</sup> Ed. Academic Press.